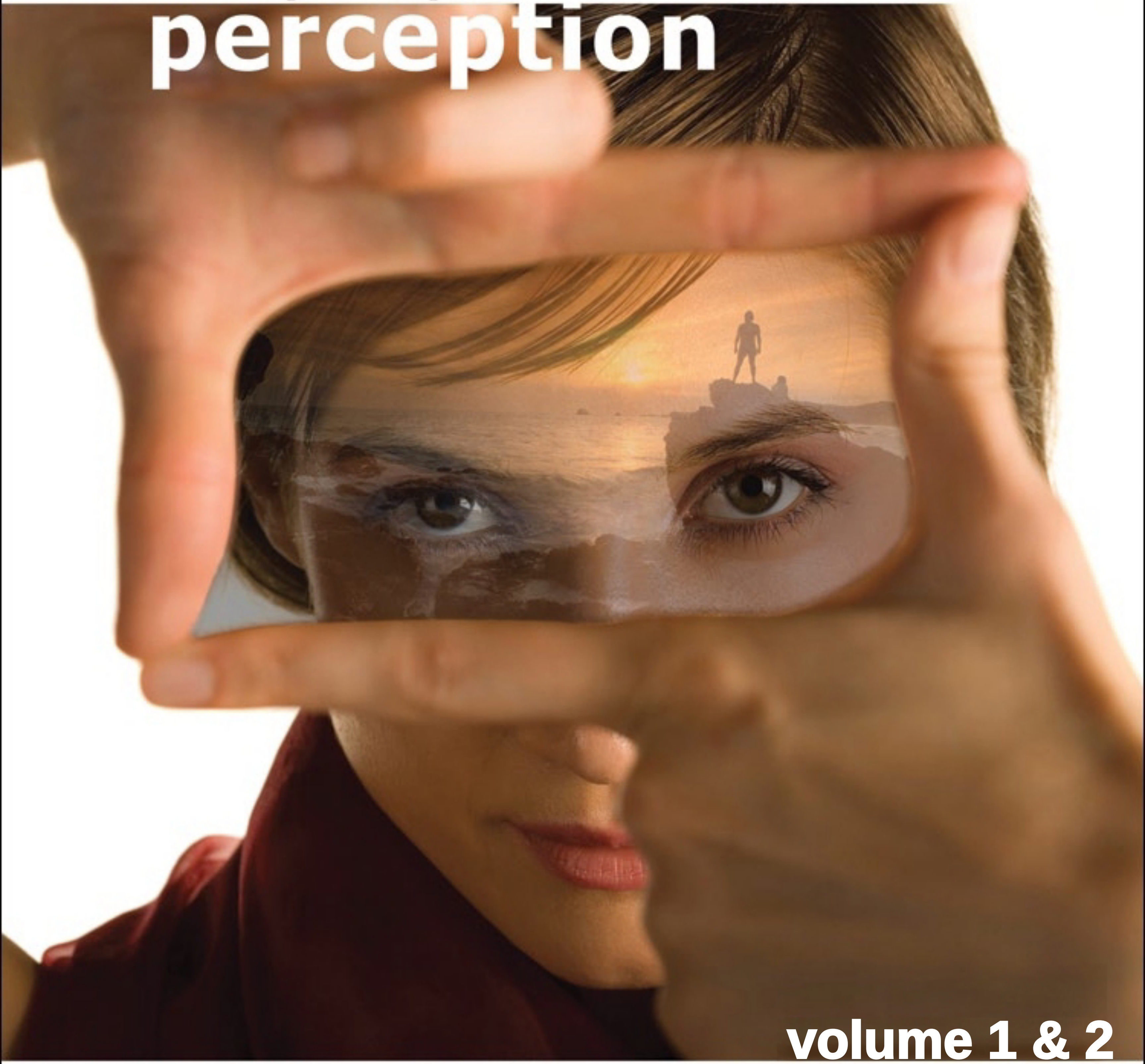
e. bruce goldstein

encyclopedia of perception



volume 1 & 2

volume 1 & 2

encyclopedia of
perception

e. bruce goldstein
university of pittsburgh
university of arizona

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

A SAGE Reference Publication

For information:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
33 Pekin Street #02-01
Far East Square
Singapore 048763

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Encyclopedia of perception/edited by E. Bruce Goldstein.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-4129-4081-8 (cloth)
1. Perception—Encyclopedias. 2. Senses and sensation—Encyclopedias. I. Goldstein, E. Bruce, 1941–

BF311.E497 2010
153.703—dc22 2009010777

This book is printed on acid-free paper.

09 10 11 12 13 10 9 8 7 6 5 4 3 2 1

Publisher:	Rolf A. Janke
Acquisitions Editor:	Jim Brace-Thompson
Editorial Assistants:	Eileen Gallaher, Michele Thompson
Developmental Editor:	Carole Maurer
Reference Systems Manager:	Leticia M. Gutierrez
Reference Systems Coordinator:	Laura Notton
Production Editor:	Kate Schroeder
Copy Editors:	Amy Freitag, Robin Gold
Typesetter:	C&M Digitals (P) Ltd.
Proofreaders:	Kevin Gleason, Annie Lubinsky
Indexer:	Virgil Diodato
Cover Designer:	Ravi Balasuriya
Marketing Manager:	Amberlyn McKay

Contents

Volume 1

Entries

Volume 2

Entries

Editor and Advisory Board

Editor

E. Bruce Goldstein
University of Pittsburgh
University of Arizona

Advisory Board

Richard N. Aslin
University of Rochester

Marlene Behrmann
Carnegie Mellon University

Randolph Blake
Vanderbilt University

Geoffrey M. Boynton
University of Washington

Paul A. S. Breslin
Monell Chemical Senses Center

Kyle R. Cave
University of Massachusetts

Melvyn Alan Goodale
University of Western Ontario

Deborah Hall
MRC Institute of Hearing Research

Mark Hollins
University of North Carolina at Chapel Hill

Tom C. T. Yin
University of Wisconsin–Madison

Steven Hsiao
Johns Hopkins University

Helene Intraub
University of Delaware

Scott P. Johnson
University of California, Los Angeles

Daniel Kersten
University of Minnesota

Roberta L. Klatzky
Carnegie Mellon University

Keith Kluender
University of Wisconsin–Madison

Lawrence E. Marks
John B. Pierce Laboratory

Stephen McAdams
McGill University

Suzanne P. McKee
Smith-Kettlewell Eye Research Institute

William A. Yost
Arizona State University

Joseph Anthony Movshon
New York University

Alva Noe
University of California, Berkeley

Christopher Plack
University of Manchester

Dennis R. Proffitt
University of Virginia

Gary B. Rollman
University of Western Ontario

Shihab A. Shamma
University of Maryland

Steven K. Shevell
University of Chicago

Donald A. Wilson
New York University
City University of New York

List of Entries

Absolute Pitch
Acoustics and Concert Halls
Action and Vision
Aesthetic Appreciation of Pictures
Aftereffects
Afterimages
Ageing and Chemical Senses
Ageing and Hearing
Ageing and Touch
Ageing and Vision
Agnosia: Auditory
Agnosia: Tactile
Agnosia: Visual
Air Quality
Amblyopia
American Sign Language
Ames Demonstrations in Perception
Amodal Perception
Animal Chemical Sensitivity
Animal Color Vision
Animal Depth Perception
Animal Eye Movements
Animal Eyes
Animal Frequency and Pitch Perception
Animal Motion Perception
Aphasias
Aromatherapy
Art and Perception. *See* Pictorial Depiction and Perception
Assistive Technologies for the Blind
Atmospheric Phenomena
Attention: Cognitive Influences
Attention: Covert
Attention: Cross-Modal
Attention: Disorders
Attention: Divided
Attention: Effect of Breakdown
Attention: Effect on Perception
Attention: Object-Based
Attention: Physiological
Attention: Selective
Attention: Spatial
Attention: Theories of
Attention and Consciousness
Attention and Emotion
Attention and Medical Diagnosis
Attention and Memory
Attractiveness
Audiology
Audition
Audition: Cognitive Influences
Audition: Disorders
Audition: Loudness
Audition: Pitch Perception
Audition: Temporal Factors
Auditory Frequency Analysis, Neural
Auditory Frequency Selectivity
Auditory Illusions
Auditory Imagery
Auditory Localization: Physiology
Auditory Localization: Psychophysics
Auditory Masking
Auditory Processing: Central
Auditory Processing: Peripheral
Auditory Receptors and Transduction
Auditory Scene Analysis
Auditory System: Damage Due to Overstimulation
Auditory System: Evolution of
Auditory System: Structure
Auditory Thresholds

Bayesian Approach
Binding Problem
Binocular Vision and Stereopsis
Bistable Perception
Body Perception
Body Perception: Disorders

Reader's Guide

The Reader's Guide is provided to assist readers in locating articles on related topics. It classifies articles into 18 general topical categories: Action; Attention; Audition; Chemical Senses; Cognition and Perception; Computers and Perception; Consciousness; Disorders of Perception; Illusory Perceptions; Individual Differences (Human) and Comparative (Across Species); Methods; Perceptual Development/Experience; Philosophical Approaches; Physiological Processes; Sense Interactions; Skin and Body Senses; Theoretical Approaches; and Visual Perception. Entries may be listed under more than one topic.

Action

Action and Vision
Corollary Discharge
Echolocation
Effort: Perception of
Embodied Perception
Event Perception
Eye and Limb Tracking
Eye Movements: Behavioral
Eye Movements: Effects of Neurological and Mental Disorders On
Eye Movements and Action in Everyday Life
Eye Movements During Cognition and Conversation
Feature Integration Theory
Film (Cinema) Perception
Guidance Systems for Blind People
Haptics
Human–Machine Interface
Kinesthesia
Mirror Neurons
Motion Parallax and Structure From Motion
Motion Perception
Motion Perception: Social
Multimodal Interactions: Visual–Haptic
Navigation Through Spatial Layout
Perceptual Development: Imitation
Perceptual Development: Touch and Pain
Perceptual Development: Visually Guided Reaching
Perceptual-Motor Integration
Prism Adaptation
Reaching and Grasping
Response Time
Self-Motion Perception
Speech Production
Tool Use
Unconscious Processes
Vestibular System
Video Games
Visually Guided Actions
Visual Search
Weight Perception

Attention

Attention: Cognitive Influences
Attention: Covert
Attention: Cross-Modal
Attention: Disorders
Attention: Divided
Attention: Effect of Breakdown
Attention: Effect on Perception
Attention: Object-Based
Attention: Physiological
Attention: Selective
Attention: Spatial
Attention: Theories of
Attention and Consciousness
Attention and Emotion
Attention and Medical Diagnosis
Attention and Memory

Bistable Perception
Cell Phones and Driver Distraction
Change Detection
Consciousness
Eye and Limb Tracking
Eye Movements During Cognition and Conversation
Film (Cinema) Perception
Magic and Perception
Perceptual Development: Attention
Rapid Serial Visual Presentation
Top-Down and Bottom-Up Processing
Video Games
Visual Search

Audition

Absolute Pitch
Acoustics and Concert Halls
Ageing and Hearing
Agnosia: Auditory
American Sign Language
Animal Frequency and Pitch Perception
Aphasias
Audiology
Audition
Audition: Cognitive Influences
Audition: Disorders
Audition: Loudness
Audition: Pitch Perception
Audition: Temporal Factors
Auditory Frequency Analysis, Neural
Auditory Frequency Selectivity
Auditory Illusions
Auditory Imagery
Auditory Localization: Physiology
Auditory Localization: Psychophysics
Auditory Masking
Auditory Processing: Central
Auditory Processing: Peripheral
Auditory Receptors and Transduction
Auditory Scene Analysis
Auditory System: Damage Due to Overstimulation
Auditory System: Evolution of
Auditory System Structure
Auditory Thresholds
Causality
Cochlear Implants: Controversy
Cochlear Implants: Technology
Computer-Generated Speech, Perception of
Computer Speech Perception
Echolocation
Evoked Potential: Audition
Guidance Systems for Blind People
Hearing Aids
Language
Lightning and Thunder
Melody Perception
Mirror Neurons
Multimodal Interactions: Tactile–Auditory
Multimodal Interactions: Visual–Auditory
Music Cognition and Perception
Music in Film
Otoacoustic Emissions
Perceptual Development: Hearing
Perceptual Development: Infant Music Perception
Perceptual Development: Intermodal Perception
Perceptual Development: Speech Perception
Sound Reproduction and Perception
Sound Stimulus
Speech Perception
Speech Perception: Physiological
Speech Production
Speechreading
Statistical Learning
Synesthesia
Timbre Perception
Tinnitus
Unconscious Processes
Virtual Reality: Auditory
Word Recognition

Chemical Senses

Ageing and the Chemical Senses
Air Quality
Animal Chemical Sensitivity
Aromatherapy
Common Chemical Sense (Chemesthesis)
Constancy
Electronic Nose
Flavor
Fragrances and Perfume
Multimodal Interactions: Color–Chemical
Multimodal Interactions: Thermal–Chemical
Olfaction
Olfaction: Evolution of
Olfaction: Feature Detection and Integration
Olfaction and Reproductive Behavior
Olfactometry
Olfactory Adaptation
Olfactory Bulb: Functional Architecture

Cognition and Perception

Computers and Perception

Consciousness

Disorders of Perception

Prostheses: Visual
Recovery of Vision
  Following Blindness
Sensory Rehabilitation
Sensory Restoration and Substitution
Speechreading
Taste: Disorders
Tinnitus
Vision: Developmental Disorders
Visual Disorders: Blindness

Illusory Perceptions

Aftereffects
Afterimages
Atmospheric Phenomena
Auditory Illusions
Change Detection
Contrast Enhancement at Borders
Contrast Perception
Hallucinations and Altered Perceptions
Impossible Figures
Nonveridical Perception
Phantom Limb
Pictorial Depiction and Perception
Visual Illusions

Individual Differences (Human) and Comparative (Across Species; Not Including Ageing, Disorders, and Perceptual Development)

Absolute Pitch
Animal Chemical Sensitivity
Animal Color Vision
Animal Depth Perception
Animal Eye Movements
Animal Eyes
Animal Frequency and Pitch Perception
Animal Motion Perception
Cultural Effects on Perception
Echolocation
Electroception
Emotional Influences on Perception
Individual Differences in Perception
Nature and Nurture in Perception
Pain: Cognitive and
  Contextual Influences
Perceptual Expertise
Private Nature of Perceptual Experience
Taste: Genetics of
Taste: Supertasters
Taste and Food Preferences
Video Games

Methods

Brain Imaging
Evoked Potential: Audition
Evoked Potential: Vision
Magnetoencephalography
Microstimulation
Neural Recording
Neuropsychology of Perception
Phenomenology (Psychology)
Physiological Approach
Priming
Psychophysical Approach
Psychophysics: Detection
Rapid Serial Visual Presentation
Receptive Fields
Recognition
Response Time
Reverse Correlation
Scaling of Sensory Magnitude
Selective Adaptation
Signal Detection Theory and Procedures
Transcranial Magnetic Stimulation
Tuning Curves
Visual Search

Perceptual Development/Experience

Ageing and Chemical Senses
Ageing and Hearing
Ageing and Touch
Ageing and Vision
Cultural Effects on Perception
Infant Perception
Infant Perception: Methods of Testing
Nature and Nurture in Perception
Perceptual Development: Attention
Perceptual Development: Color and Contrast
Perceptual Development: Face Perception
Perceptual Development: Hearing
Perceptual Development: Imitation
Perceptual Development: Intermodal Perception
Perceptual Development: Object Perception
Perceptual Development: Taste and Olfaction
Perceptual Development: Touch and Pain
Perceptual Development: Visual Acuity
Perceptual Development: Visually Guided Reaching
Perceptual Development: Visual Object
  Permanence and Identity

Perceptual Expertise
Perceptual Learning
Prism Adaptation
Statistical Learning
Vision: Developmental Disorders

Philosophical Approaches

Causality
Color: Philosophical Issues
Computer Consciousness
Consciousness
Content of Perceptual Experience
Indirect Nature of Perception
Intentionality and Perception
Inverted Spectrum
Mary the Color Scientist
Mind and Body
Modality (Philosophy)
Molyneux's Question
Naïve Realism
Perceptual Representation (Philosophy)
Phenomenology (Philosophy)
Philosophical Approaches
Philosophy: Access and Report
Philosophy: Attention and the Size of the Conscious Field
Private Nature of Perceptual Experience
Qualia
Seeing As
Visual Filling In and Completion

Physiological Processes

Aftereffects
Binding Problem
Contrast Enhancement at Borders
Cortical Organization
Cortical Reorganization Following Damage
Experience-Dependent Plasticity
Feedback Pathways
Lateral Inhibition
Loss of a Sense, Effect on Others, Psychological
Mirror Neurons
Modularity
Multimodal Interactions: Neural Basis
Neural Prosthetic Systems
Neural Recording
Neural Representation/Coding
Neuropsychology of Perception
Oscillatory Synchrony
Physiological Approach
Receptive Fields
Speed of Processing in Sensory Systems
Tuning Curves

Sense Interactions

Action and Vision
Attention: Cross-Modal
Cortical Reorganization Following Damage
Cross-Modal Transfer
Extrasensory Perception
Flavor
Loss of a Sense: Effect on Others, Psychological
Molyneux's Question
Motion Perception: Social
Multimodal Interactions: Color–Chemical
Multimodal Interactions: Neural Basis
Multimodal Interactions: Pain–Touch
Multimodal Interactions: Tactile–Auditory
Multimodal Interactions: Thermal–Chemical
Multimodal Interactions: Visual–Auditory
Multimodal Interactions: Visual–Haptic
Perceptual Development: Intermodal Perception
Perceptual-Motor Integration
Sensory Restoration and Substitution
Synesthesia
Taste and Food Preferences

Skin and Body Senses

Ageing and Touch
Agnosia: Tactile
Body Perception
Body Perception: Disorders
Braille
Constancy
Cutaneous Perception
Cutaneous Perception: Physiology
Electroception
Embodied Perception
Haptics
Itch, Tickle, and Tingle
Kinesthesia
Migraine
Molyneux's Question
Multimodal Interactions: Pain–Touch
Multimodal Interactions: Tactile–Auditory
Multimodal Interactions: Thermal–Chemical
Multimodal Interactions: Visual–Haptic

Out-of-Body Experience
Pain: Assessment and Measurement
Pain: Cognitive and Contextual Influences
Pain: Neuromatrix Theory
Pain: Physiological Mechanisms
Pain: Placebo Effects
Pain: Treatments for Chronic
Perceptual Development: Touch and Pain
Phantom Limb
Proprioception
Reaching and Grasping
Surface and Material Properties Perception
Tactile Acuity
Tactile Map Reading
Temperature Perception
Texture Perception: Tactile
Tool Use
Vibratory Perception
Virtual Reality: Touch/Haptics
Visceral Perception
Weight Perception

Theoretical Approaches

Bayesian Approach
Computational Approaches
Direct Perception
Ecological Approach
Embodied Perception
Evolutionary Approach
Evolutionary Approach: Perceptual Adaptations
Gestalt Approach
Indirect Nature of Perception
Information Theory
Physiological Approach
Psychophysical Approach
Theoretical Approaches
Theory of Mind

Visual Perception

Action and Vision
Aesthetic Appreciation of Pictures
Aftereffects
Afterimages
Ageing and Vision
Agnosia: Visual
Amblyopia
American Sign Language
Ames Demonstrations in Perception
Amodal Perception
Animal Color Vision
Animal Depth Perception
Animal Eye Movements
Animal Eyes
Animal Motion Perception
Assistive Technologies for the Blind
Atmospheric Phenomena
Attention: Cognitive Influences
Attention: Covert
Attention: Cross-Modal
Attention: Disorders
Attention: Divided
Attention: Effect of Breakdown
Attention: Effect on Perception
Attention: Object-Based
Attention: Physiological
Attention: Selective
Attention: Spatial
Attention: Theories of
Attention and Consciousness
Attention and Emotion
Attention and Medical Diagnosis
Attractiveness
Binding Problem
Binocular Vision and Stereopsis
Bistable Perception
Camouflage
Causality
Change Detection
Color: Genetics of
Color: Philosophical Issues
Color Constancy
Color Deficiency
Color Mixing
Color Naming
Color Perception
Color Perception: Physiological
Computer Graphics and Perception
Computer Vision
Constancy
Context Effects in Perception
Contrast Perception
Corollary Discharge
Depth Perception in Pictures/Film
Digital Imaging
Direct Perception
Dyslexia
Ecological Approach
Embodied Perception
Event Perception

About the Editor

E. Bruce Goldstein is Professor Emeritus of Psychology at the University of Pittsburgh and Adjunct Professor of Psychology at the University of Arizona. He has received the Chancellor's Distinguished Teaching Award from the University of Pittsburgh for his classroom teaching and textbook writing. He received his bachelor's degree in chemical engineering from Tufts University, his PhD in experimental psychology from Brown University, and he was a post-doctoral fellow in the Biology Department at Harvard University before joining the faculty at the University of Pittsburgh. Dr. Goldstein has published papers on a wide variety of topics, including retinal and cortical physiology, visual attention, and the perception of pictures. He is the author of *Sensation & Perception,* 8th edition (Wadsworth/Cengage, 2010), and *Cognitive Psychology: Connecting Mind, Research and Everyday Experience,* 2nd edition (Wadsworth/Cengage, 2008), and is the editor of the *Blackwell Handbook of Perception* (Blackwell, 2001).

Contributors

Mark Addis
Birmingham City University

Michael A. Akeroyd
MRC Institute of Hearing Research

Mark V. Albert
Cornell University

Robert S. Allison
York University

Torin Alter
University of Alabama

Dima Amso
Weill Medical College of Cornell University

Ulrich Ansorge
University of Osnabrueck

L. Gregory Appelbaum
Duke University

Ian A. Apperly
University of Birmingham

Aries Arditi
Lighthouse International

Pablo Artal
Universidad de Murcia

F. Gregory Ashby
University of California, Santa Barbara

Janette Atkinson
University College London

Michael Bach
University of Freiburg

David Baguley
Cambridge University Hospital

Lorraine E. Bahrick
Florida International University

Jason J. S. Barton
University of British Columbia

Daphne Bavelier
University of Rochester

Roy R. Behrens
University of Northern Iowa

Marlene Behrmann
Carnegie Mellon University

Fabrizio Benedetti
University of Turin Medical School

Sliman Bensmaia
Johns Hopkins University

Tonya R. Bergeson
Indiana University School of Medicine

Lynne Esther Bernstein
House Ear Institute

Geoffrey P. Bingham
Indiana University

Olaf Blanke
Ecole Polytechnique Federale de Lausanne (EPFL)

John D. Bonvillian
University of Virginia

Richard T. Born
Harvard Medical School

John Brabyn
Smith-Kettlewell Institute

Oliver Braddick
University of Oxford

David H. Brainard
University of Pennsylvania

Myron L. Braunstein
University of California, Irvine

Mary Bravo
Rutgers University

Peter A. Brennan
University of Bristol

Paul A. S. Breslin
Monell Chemical Senses Center

Farran Briggs
University of California, Davis

James R. Brockmole
University of Edinburgh

Peter Brugger
University Hospital Zurich

Isabelle Bülthoff
Max Planck Institute of Biological Cybernetics

Dwight A. Burkhardt
University of Minnesota

Elke K. Buschbeck
University of Cincinnati

Laurel J. Buxbaum
Thomas Jefferson University

Patrick Byrne
York University

Gianluca Campana
University of Padova

Laura A. Carlson
University of Notre Dame

Catherine E. Carr
University of Maryland

Marisa Carrasco
New York University

Joseph Carroll
Medical College of Wisconsin

E. Carstens
University of California, Davis

Vivien A. Casagrande
Vanderbilt Medical School

Craig A. Champlin
University of Texas at Austin

Roger W. Cholewiak
Princeton University

Steven A. Cholewiak
Rutgers University

Jakob Christensen-Dalsgaard
University of Southern Denmark

Charles Chubb
University of California, Irvine

Susana T. L. Chung
University of Houston

Jennifer Church
Vassar College

Thomas A. Cleland
Cornell University

Colin W. G. Clifford
University of Sydney

David Cole
University of Minnesota, Duluth

Han Collewijn
Erasmus University

J. Enrique Cometto-Muniz
University of California, San Diego

Christopher M. Conway
Indiana University

Erik P. Cook
McGill University

Joel M. Cooper
University of Utah

Daniela Corbetta
University of Tennessee

Lawrence K. Cormack
University of Texas

Brian D. Corneil
University of Western Ontario

Beverly J. Cowart
Monell Chemical Senses Center

Trinity B. Crapse
University of Pittsburgh

John Douglas Crawford
York University

Jody C. Culham
University of Western Ontario

Delphine Dahan
University of Pennsylvania

James L. Dannemiller
Rice University

Chris Darwin
University of Sussex

Charles Derby
Georgia State University

Amy S. Desroches
University of Western Ontario

Diana Deutsch
University of California, San Diego

P. De Weerd
Maastricht University, The Netherlands

Vincent Di Lollo
Simon Fraser University

Patricia M. Di Lorenzo
Binghamton University

Hubert Dinse
Ruhr-University Bochum

Jochen Ditterich
University of California, Davis

Tim Donovan
University of Cumbria

Barbara Dosher
University of California, Irvine

Richard L. Doty
University of Pennsylvania School of Medicine

W. Jay Dowling
University of Texas at Dallas

Judy R. Dubno
Medical University of South Carolina

Grant E. DuBois
Coca-Cola Company

Frank H. Durgin
Swarthmore College

Matthew W. G. Dye
University of Rochester

Ben Dyson
Ryerson University

Ruth Anne Eatock
Harvard Medical School

R. A. R. Edwards
Rochester Institute of Technology

Jos J. Eggermont
University of Calgary

Heather L. Eisthen
Michigan State University

William Epstein
University of Virginia

Hany Farid
Dartmouth College

Olga Eizner Favreau
University of Montreal

Jeremy D. Fesi
Pennsylvania State University

David J. Field
Cornell University

Tiffany Field
University of Miami

Ione Fine
University of Washington

Max L. Fletcher
Yale University School of Medicine

Jonathan Isaac Flombaum
Johns Hopkins University

Mary Florentine
Northeastern University

John M. Foley
University of California, Santa Barbara

Charles L. Folk
Villanova University

Catherine A. Forestell
College of William & Mary

David J. Freedman
University of Chicago

Barrie Frost
Queens University

Isabel Gauthier
Vanderbilt University

Karl Gegenfurtner
Justus-Liebig-Universität

Michael E. Geisser
University of Michigan

Mark Georgeson
Aston University

Alan Gilchrist
Rutgers University

Rick O. Gilmore
Pennsylvania State University

Jason M. Gold
Indiana University

E. Bruce Goldstein
University of Pittsburgh, University of Arizona

Phil Gomersall
Cambridge University Foundation Hospitals Trust

Melvyn Alan Goodale
University of Western Ontario

Richard H. Gracely
University of Michigan Health System

C. S. Green
University of Rochester

Timothy D. Griffiths
Newcastle University

Kalanit Grill-Spector
Stanford University

Simon Grondin
Université Laval

Ricardo Gutierrez-Osuna
Texas A&M University

Carole M. Hackney
University of Cambridge

Deborah Hall
MRC Institute of Hearing Research

Thorsten Hansen
Justus-Liebig-University Giessen

C. L. Hardin
Syracuse University

Alison M. Harris
University of Pennsylvania

E. Eugenie Hartmann
University of Alabama at Birmingham

Mary Hayhoe
University of Texas at Austin

Marcia J. Hay-McCutcheon
Indiana University School of Medicine

Sheng He
University of Minnesota

Heiko H. Hecht
Johannes Gutenberg-Universität Mainz

David J. Heeger
New York University

Harry Heft
Denison University

Mary Hegarty
University of California, Santa Barbara

Christine Heinisch
Ruhr-University

Michael G. Heinz
Purdue University

Morton A. Heller
Eastern Illinois University

Rachel S. Herz
Brown University

David Hilbert
University of Illinois at Chicago

Donald D. Hoffman
University of California, Irvine

Alex O. Holcombe
University of Sydney

George Hollich
Purdue University

Mark Hollins
University of North Carolina at Chapel Hill

Carl D. Hopkins
Cornell University

Todd S. Horowitz
Harvard Medical School

Derek M. Houston
Indiana University School of Medicine

Steven Hsiao
Johns Hopkins University

Ludwig Huber
University of Vienna

Thomas Hummel
University of Dresden Medical School

Giuseppe Iaria
University of British Columbia

Ronald S. Jackson
CCOVI, Brock University

Gerald H. Jacobs
University of California, Santa Barbara

R. Daniel Jacobson
University of Calgary

Marc F. Joanisse
University of Western Ontario

Kerri Johnson
University of California, Los Angeles

Scott P. Johnson
University of California, Los Angeles

Lynette A. Jones
Massachusetts Institute of Technology

J. Scott Jordan
Illinois State University

Jon Kaas
Vanderbilt University

Jay Kadis
Stanford University

Tobias Kalisch
Ruhr-University

Jian Kang
University of Sheffield

Joel Katz
York University

Brian L. Keeley
Pitzer College

Vladimir J. Kefalov
Washington University School of Medicine

Jonathan W. Kelly
Vanderbilt University

Roger A. Kendall
University of California, Los Angeles

Dirk Kerzel
Université de Genève

Mudasar Khan
Cleveland Police Services

Gerald Kidd, Jr.
Boston University

HoJin Kim
University of California, Los Angeles

Roberta L. Klatzky
Carnegie Mellon University

Stanley A. Klein
University of California, Berkeley

Eliana M. Klier
Washington University School of Medicine

Guenther Knoblich
Radboud University Nijmegen

Zoe Kourtzi
University of Birmingham

Eileen Kowler
Rutgers University

Richard J. Krauzlis
Salk Institute for Biological Studies

Gustav Kuhn
Durham University

James R. Lackner
Brandeis University

Michael S. Landy
New York University

Susan J. Lederman
Queen's University

Barry B. Lee
State University of New York College of Optometry

Hyunkyu Lee
University of Illinois

Kang Lee
University of Toronto

Tai Sing Lee
Carnegie Mellon University

Gordon E. Legge
University of Minnesota

Christian H. Lemon
St. Louis University School of Medicine

Bigna Lenggenhager
Ecole Polytechnique Federale de Lausanne (EPFL)

Dennis M. Levi
University of California, Berkeley

Daniel Levitin
McGill University

Jonathan B. Levitt
City College of New York

David J. Lewkowicz
Florida Atlantic University

Robert Lickliter
Florida International University

Jeremy L. Loebach
Indiana University

Elizabeth F. Loftus
University of California, Irvine

Jack M. Loomis
University of California, Santa Barbara

Sanford Lopater
Christopher Newport University

Zhong-Lin Lu
University of Southern California

Gary Lupyan
Cornell University

Robert A. Lutfi
University of Wisconsin–Madison

Wei Ji Ma
Baylor College of Medicine

William Mace
Trinity College

Arien Mack
New School for Social Research

Stephen L. Macknik
Barrow Neurological Institute

Peter F. MacNeilage
University of Texas at Austin

Fiona Macpherson
University of Glasgow

Walter Makous
University of Rochester

Jitendra Malik
University of California, Berkeley

Laurence T. Maloney
New York University

Samuele Marcora
Bangor University

William L. Martens
University of Sydney

Susana L. Martinez-Conde
Barrow Neurological Institute

Takahiko Masuda
University of Alberta

Mohan Matthen
University of Toronto

Daphne Maurer
McMaster University

David McAlpine
University College London

Donald H. McBurney
University of Pittsburgh

John J. McCann
McCann Imaging

Martha K. McClintock
University of Chicago

Robert D. McIntosh
University of Edinburgh

Timothy P. McNamara
Vanderbilt University

Robert M. McPeek
Smith-Kettlewell Eye Research Institute

Sara C. Mednick
University of California, San Diego

Andrew N. Meltzoff
University of Washington

Ronald Melzack
McGill University

Janine D. Mendola
McGill University

Susanna Millar
University of Oxford

Lee M. Miller
University of California, Davis

Partha Pratim Mitra
Cold Spring Harbor Laboratory

Richard T. Miyamoto
Indiana University School of Medicine

Brian C. J. Moore
University of Cambridge

Cathleen M. Moore
University of Iowa

Michael Morgan
City University London

Cynthia F. Moss
University of Maryland

Steven B. Most
University of Delaware

Steven D. Munger
University of Maryland School of Medicine

Kevin J. Munro
University of Manchester

Sam Musallam
McGill University

Bence Nanay
Syracuse University

Amy Needham
Duke University

Matthew Nudds
University of Edinburgh

Howard C. Nusbaum
University of Chicago

Robert P. O'Shea
University of Otago

Aude Oliva
Massachusetts Institute of Technology

Lynn A. Olzak
Miami University of Ohio

Yuri Ostrovsky
Massachusetts Institute of Technology

Andrew Oxenham
University of Minnesota

Anna Papafragou
University of Delaware

Andrew Richard Parker
Green College, Oxford

Olivier Pascalis
University of Sheffield

Kevin A. Pelphrey
Yale University

James W. Pennebaker
University of Texas at Austin

John Philbeck
George Washington University

David B. Pisoni
Indiana University

Alexander Pollatsek
University of Massachusetts, Amherst

James R. Pomerantz
Rice University

Ariella Popple
University of California, Berkeley

Mary C. Portillo
Rice University

J. Prescott
University of Newcastle

Gopathy Purushothaman
Vanderbilt Medical School

Philip Quinlan
University of York

Paul C. Quinn
University of Delaware

Josef P. Rauschecker
Georgetown University Medical Center

Keith Rayner
University of California, San Diego

Gregg H. Recanzone
University of California, Davis

Gordon M. Redding
Illinois State University

Catherine L. Reed
Claremont McKenna College, University of Denver

Geraint Rees
University College London

Adam J. Reeves
Northeastern University

Daniel Reisberg
Reed College

Ronald A. Rensink
University of British Columbia

Daniel C. Richardson
University College London

Lynn C. Robertson
University of California, Berkeley

Gabriel Robles-De-La-Torre
International Society for Haptics

Urte Roeber
University of Leipzig

Sheena Rogers
James Madison University

Gary B. Rollman
University of Western Ontario

Andrew M. Rosen
Binghamton University

Helen E. Ross
University of Stirling

Richard Russell
Harvard University

Navid G. Sadeghi
McGill University

Susan S. Schiffman
Duke University Medical Center

James A. Schirillo
Wake Forest University

Brian J. Scholl
Yale University

Martin Schürmann
University of Nottingham

Eric Schwitzgebel
University of California, Riverside

Natalie Sebanz
Radboud University Nijmegen

H. A. Sedgwick
State University of New York College of Optometry

Mark A. Segraves
Northwestern University

Allison B. Sekuler
McMaster University

Muhammad Shah
University of Pennsylvania

Ladan Shams
University of California, Los Angeles

Stanley Sheft
Loyola University Chicago

Peter Shirley
University of Utah

William P. Shofner
Indiana University

Sarah Shomstein
George Washington University

Susanna Siegel
Harvard University

Wayne L. Silver
Wake Forest University

Christopher Simons
Givaudan Flavors Corporation

Donal G. Sinex
Utah State University

Pawan Sinha
Massachusetts Institute of Technology

Ruxandra Sireteanu
Max-Planck-Institute for Brain Research

James A. Skelton
Dickinson College

Jay Slack
Givaudan Flavors Corporation

Alan M. Slater
University of Exeter

Monique A. M. Smeets
Utrecht University

Tim J. Smith
University of Edinburgh

Marc A. Sommer
University of Pittsburgh

Ferrinne Spector
McMaster University

Charles Spence
Oxford University

Nancy K. Steblay
Augsburg College

Richard J. Stevenson
Macquarie University

Andrew Stockman
University College London

David L. Strayer
University of Utah

Veit Stuphorn
Johns Hopkins University

Diane M. Tait
Medical College of Wisconsin

Catherine Tallon-Baudry
Centre national de la recherche scientifique

Desney S. Tan
Microsoft Research

Martin Tegenthoff
Ruhr-University

Hiroko Teresawa
Stanford University

William B. Thompson
University of Utah

Simon J. Thorpe
Centre national de la recherche scientifique

Brian Timney
University of Western Ontario

Anna K. Tirovolas
McGill University

Daniel J. Tollin
University of Colorado Health Sciences Center

Connie A. Tompkins
University of Pittsburgh

Laurel J. Trainor
McMaster University

Anne Treisman
Princeton University

Hely Tuorila
University of Helsinki

Nicholas B. Turk-Browne
Yale University

W. Martin Usrey
University of California, Davis

Michael Valente
Washington University School of Medicine

Kenneth F. Valyear
University of Western Ontario

Shaun P. Vecera
University of Iowa

Quoc C. Vuong
Newcastle University

Mark N. Wallace
MRC Institute of Hearing Research, Nottingham

Mark T. Wallace
Vanderbilt University

Vincent Walsh
University College London

Jamie Ward
University of Sussex

Lawrence M. Ward
University of British Columbia

William H. Warren
Brown University

Michael Webster
University of Nevada, Reno

Lynne A. Werner
University of Washington

David A. Westwood
Dalhousie University

Theresa L. White
Le Moyne College

Laurie M. Wilcox
York University

Teresa Wilcox
Texas A&M University

Anna Wilkinson
University of Vienna

Frances Wilkinson
York University

Donald A. Wilson
New York University, City University of New York

Hugh R. Wilson
York University

Paul Wise
Monell Chemical Senses Center

William A. Yost
Arizona State University

Debra A. Zellner
Montclair State University

Introduction

The Field of Perception

The field of perception is concerned with explaining the operation of the senses and the experiences and behaviors resulting from stimulation of the senses. The senses are vision, hearing, the cutaneous senses (touch, pain, tickle, itch), chemical senses (taste, smell, flavor), proprioception and kinesthesia (awareness of body positions and limb position and motion), and the vestibular sense (balance).

One of the characteristics of perception is that it involves more than is apparent. This statement reflects the following paradox: Perceptual experiences such as recognizing objects; seeing light, color, and spatial layout; hearing pitch and music; perceiving the locations of sounds; and experiencing smells, tastes, touch, cold, and pain represent the actions of extremely complex, and still not completely understood, mechanisms. Yet for all the underlying complexity of these mechanisms, we perceive easily, usually with little effort or conscious awareness of the mechanisms involved. One way to appreciate the fact that there is more to perception than may be apparent to the perceiver is to consider the following examples.

- A person looks out over a visual scene dotted with houses, trees, and other objects. Although this scene creates a two-dimensional image on the person's retina, the person sees the scene as stretching into the distance, and the objects as volumetric shapes.
- Bacon simmering and coffee brewing release hundreds of different types of molecules into the air. The molecules from the bacon and coffee become indiscriminately mixed in the air and across a person's olfactory receptors. Despite this random mixing of molecules, the person perceives just two olfactory objects, "bacon" and "coffee." This feat is particularly impressive because many of the individual molecules, when considered individually, have their own odors, which do not necessarily resemble the odors of bacon or coffee.
- It is very easy for people to perceive and recognize objects in a scene ("that's a chair in the corner," "that's Sandra's face"). However, the most powerful computers can accomplish this task only with difficulty and cannot begin to approach the speed and accuracy of human performance, especially under "real-world" conditions when objects are seen at different distances and from different angles, and are sometimes partially obscured by other objects.
- A woman's arm was amputated after she was injured in a car accident. Although her arm and hand are no longer there, she still feels as if they are. In fact, she sometimes has the disconcerting experience of feeling the missing hand as tightly clenched, with fingers digging painfully into her palm.
- All of these examples and everything else we perceive are created by electrical signals in the brain, which in themselves have no color, sound, taste, hot, or cold. Yet somehow these electrical signals become transformed into perceptual experience.

This encyclopedia provides an overview of the field of perception through authoritative 1,000- to 4,000-word essays by leading researchers and theoreticians in the field. These essays are supported by nearly 180 figures and tables, 39 of which are in color.

The research and theory reported here involves two parallel and interacting approaches, the psychophysical approach and the physiological approach. The psychophysical approach involves determining the relationship between stimuli in the environment and perception. This approach has measured basic operating characteristics of perception, which involves determining things such as thresholds for qualities associated with each sense (for example, detecting light, motion, sounds, taste, and smell stimuli) and how stimulus characteristics influence perception of these qualities (for example, how wavelengths of light are associated with color perception, frequency of sound with pitch, chemical structures with taste and smell). This approach is also important for uncovering underlying mechanisms of perception and has provided the basis of various theoretical approaches to perception.

The physiological approach is concerned with determining the biological mechanisms responsible for perception. These physiological mechanisms begin when receptors sensitive to specific kinds of stimuli (light for vision, sound for hearing, for example) are activated and trigger electrical signals that eventually result in activation of neurons in a large number of structures, both before reaching the brain and then within the complex architecture of a large number of interrelated brain areas.

The physiological approach has determined how stimuli in the environment are represented by the firing of neurons and activation of the brain, and has demonstrated connections between specific areas of the brain and the perception of specific types of stimuli (for example, visual patterns such as faces, complex sounds, and chemical compounds with specific structural components have been linked to specific brain areas). The complexity of the brain processes that cause perception is reflected by the finding that although brain areas have been identified that process information about specific types of stimuli, these stimuli also cause activity in many other areas of the brain, as well.

Perception is not, however, determined only by receptors sending signals to the brain. Other factors such as the context within which perception is experienced, individual differences across perceivers, and a person's expectations and prior knowledge can also influence what is perceived. In addition, "perceptual experience"—seeing a tree, smelling the scent of a rose, feeling pain from an injury—is not the only outcome of stimulating sensory receptors. Once an object is perceived, it is usually "recognized," by placing it in a category ("that's a cup," "that's the smell of a rose"), and then in some cases the person interacts with stimuli in some way. This can involve using perceptual landmarks to help navigate through the environment, deciding to pick up a coffee cup, or accepting or rejecting a particular food based on its taste or smell.

Thus, although many of the entries in these volumes are about sensing stimuli, many other entries illustrate how the field of perception extends to "recognizing," "taking action," and other areas as well. Perception is, for example, linked to processes such as memory (memory can be enhanced for experiences rich in perceptual detail; perception can be influenced by past experiences), thinking (perception can both aid thinking and involves processes similar to those involved in thinking), emotion (perceptual experiences can create both positive and negative emotions; emotional states can influence perceptual behavior), and motivation (perceptions such as smell and pain are associated with approach and avoidance behavior; a person's goals and intentions guide looking behavior and tactile exploration).

Are there processes that are *not* associated with or influenced by perception? Surely there are—but the pervasiveness of perception is truly impressive, and the phenomena of perception and the mechanisms underlying these phenomena are what this encyclopedia is about.

Goals of the Encyclopedia

Three basic goals guided the creation of this encyclopedia: (1) broad coverage, (2) authoritative entries, and (3) accessibility to a general audience. The brief overview of the field at the beginning of this introduction describes the field of perception as involving (a) many different sensory qualities and physiological systems; (b) two broad approaches, each of which makes use of a wide variety of methods; and (c) links to many areas outside the field. Thus, in deciding how to represent the field of perception, my guiding principle was to include as varied and broad a list of topics as possible, and to include not only the basic research that accounts for most of the entries, but also descriptions of methods, theoretical approaches, and real-world applications of perceptual research. Although the primary emphasis is on explaining human perception, much animal research is included, both because of its importance in its own right and because of what the results of this research tell us about human perception.

Given the goal of casting a broad net, it is not surprising that researchers represented in this volume come from a wide range of backgrounds. For example, the list of contributors beginning on page xix includes people in departments of anatomy, behavioral science, biology, biophysics, cognitive science, computer science, neuroscience, ophthalmology, optometry, philosophy, physiology, psychology, statistics, and zoology.

The second goal, presenting authoritative state-of-the-art information, has been achieved by inviting recognized experts to write the entries. But having experts write entries is useful only to the extent that these entries are accessible to our readers. Thus, the third goal was to be sure that the experts wrote entries that could be understand by the target audience for encyclopedia—general readers with no previous background in perception, which includes anyone with an interest in perception—undergraduates, college graduates, and members of the general public. But even though entries have been written with the goal of accessibility to a wide audience, there is information here for specialists as well. Even though I have a broad background in perception, I learned quite a bit in the process of reading these entries.

Content and Organization

As indicated, the content of this encyclopedia includes topics that span the field of perception. This scope can be appreciated by consulting the List of Entries beginning on page v, which lists the 367 headwords in the alphabetical order in which they appear in the encyclopedia. Another resource is the Reader's Guide, beginning on page xi, which includes the following 18 headings with relevant headwords listed under each heading:

Action
Attention
Audition
Chemical Senses
Cognition and Perception
Computers and Perception
Consciousness
Disorders of Perception
Illusory Perceptions
Individual Differences (Human) and Comparative (Across Species; Not Including Ageing, Disorders, and Perceptual Development)
Methods
Perceptual Development/Experience
Philosophical Approaches
Physiological Processes
Sense Interactions
Skin and Body Senses
Theoretical Approaches
Visual Perception

Although headwords appear alphabetically in the encyclopedia (that is, they are not grouped by the headings listed), these headings serve as a guide to the topics that are relevant to each heading. The fact that a particular topic may appear under a number of headings reflects the interrelatedness between areas of perception. Additional evidence for this interrelatedness is the list of related topics ("*See also*" cross-references) that appears at the end of each entry. Also, Further Readings at the end of each entry provide a starting point for further, more detailed study of a topic.

One of the key features of the encyclopedia is the 16-page section of color illustrations and photos at the front of each volume, beginning on page xxxiii. Each is linked to specific entries and, at the appropriate points in the text, the entries send the reader to examine the figures. The interrelationship between the art and entries helps the encyclopedia to be useful for all readers.

Finally, a word about the headwords. Headwords were assigned lengths from 1,000 to 4,000 words based on the generality or broadness of the topics, with more general topics being assigned more words. However, even many of the shortest entries could be expanded to fill a book (and in many cases, our contributors are authors of such books).

A few of the longest entries are intended to provide an overview of an area, so the authors of these entries were asked to provide perspective regarding the history and scope of the topic and not to be concerned with details that would be covered by other entries. Some of these "overview" entries are

Audition
Consciousness
Consciousness: Disorders
Cortical Organization

Cutaneous Perception
Infant Perception
Nature and Nurture in Perception
Neuropsychology of Perception
Olfaction
Philosophical Approaches
Physiological Approach
Psychophysical Approach
Taste
Theoretical Approaches
Vision

How the Encyclopedia Was Created

The creation of this encyclopedia, which spanned two years from initial conception to publication, followed these steps:

1. I invited internationally recognized experts in the field to serve on the advisory board, with the understanding that their main task would be to play the crucial role of suggesting headwords and possible contributors to write the essays for the headwords.

2. I created a list of headwords by consulting textbooks, professional journals, handbooks, and other encyclopedias (although it is important to note that there was no encyclopedia of perception to consult because one has not existed until now). "Scope" statements were also created for each headword to indicate the intended content.

3. The advisory board reviewed the headword-scope list, suggesting additions and deletions and making comments regarding the content for each entry.

4. The advisory board suggested possible contributors for the headwords in their areas of expertise. A list of potential contributors was created based on these suggestions (giving preference to people getting more "votes" from the advisory board).

5. Invitations were sent to potential contributors for each headword. When the "first choice" was unavailable, usually because of time issues, additional people were contacted until there was a contributor for each headword.

6. I read all of the entries (some in collaboration with Carole Maurer, Sage's developmental editor for the encyclopedia) and returned them for revision when necessary. Revision was requested for two reasons: (1) need for additional coverage, or in some cases less, if there was too much overlap with another entry; (2) need for improved readability. Because accessibility of the entries was one of the major goals of the encyclopedia, some entries needed to be simplified from what would appear in the research literature, and one of the main reasons for requesting revision was the need to define technical terms. In addition, the use of concrete examples was encouraged. In a few cases, entries were sent to a member of the advisory board for a second opinion.

Using the Encyclopedia

This encyclopedia can be used both to answer specific questions about perception and to obtain an overview of the field. Two excellent starting points are the List of Entries and the listing of topics in the Reader's Guide. It is often useful to check an entry's related topics ("*See also*" cross-references) for more breadth and the Further Readings for additional information and sources.

If, as is likely, you were drawn to the encyclopedia in searching for information about a specific topic, I invite you to extend your search beyond that topic. Take some time to browse. You may be surprised by unexpected revelations about specific topics and might gain some insight into the vast scope of the field of perception. After all, perceiving is something we are doing constantly, and understanding perception can enhance this experience. I've found that studying perception has made me more observant of my environment, and more appreciative of the miraculous process that transforms energy falling on receptors into the richness of experience. I hope that reading selections from this book helps you appreciate both the complexity and the beauty of the mechanisms responsible for perception.

E. Bruce Goldstein
University of Pittsburgh
University of Arizona

Acknowledgments

I am especially indebted to three people. I thank James Brace-Thompson (whom I have always known as JBT) for broaching the idea that I edit this encyclopedia, for being willing to wait until I had time to do it, and for providing the support I needed to be able to handle the logistics involved in this rather huge undertaking.

The major support JBT provided were the two people without whom I simply wouldn't have been able to make this book happen. Eileen Gallaher, my editorial assistant, sent out invitations and follow-ups, answered contributor's questions, and kept things organized (not a trivial task!) as the encyclopedia took shape. I worked with her for more than a year, sometimes on almost a daily basis, and she always came through. Thank you, Eileen!

The other person, Carole Maurer, Sage's developmental editor for the encyclopedia, helped evaluate entries, gave me advice, made revisions when necessary, and perhaps most important, provided emotional support. Working with Carole, which felt like a true collaboration, was a pleasure! Thank you, Carole.

Once all the entries were in and Carole had transformed them into the correct format, she handed the manuscript off to Kate Schroeder, the production editor, who assigned Amy Freitag and Robin Gold to copyedit the manuscript. Thank you Amy and Robin for your excellent copyediting. Also a special thanks to Kate for everything you did during the challenging process of transforming the manuscript into this book.

There are two groups of people, without whom this book wouldn't exist—my advisory board and the contributors. I thank the members of the advisory board, whose names are listed on page iv. Their major job was to help get the project off the ground by suggesting headwords, scope statements, and contributors. Only with their help was I able to cast the net out into the field of perception that made it possible for me to recruit the distinguished list of contributors who wrote the 367 entries. I especially thank Mark Hollins, Don Wilson, and Bill Yost, who provided sample entries to post on the Web site, and a special thanks to Don Wilson for being willing to provide advice on an ongoing basis as the project progressed.

And finally, I thank the hundreds of contributors, whose names are listed beginning on page xix. Everyone who wrote an entry was extremely busy with other projects but somehow made time to write her or her entries. But what impressed me the most about the contributors was the graciousness with which they responded to my requests for revision. Much to my surprise, some of them even thanked me for my suggestions! In addition, I had some interesting and intellectually stimulating e-mail conversations with a number of the contributors. Thank you all for making the journey from agreeing to edit the encyclopedia to seeing the completed book a positive one for me, and for making the final product useful to our readers, and, I hope also to the field of perception.

Bruce Goldstein

List of Color Illustrations

Audition

Cutaneous

Olfaction

Vision

General Perception

Audition

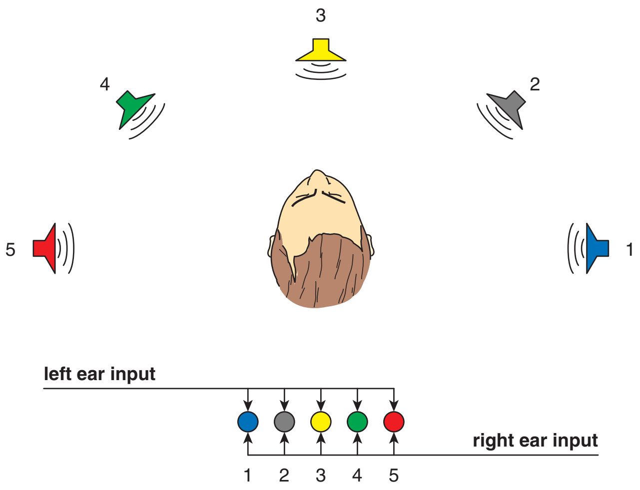


Figure 1 *Auditory Localization: Physiology*—Axonal Delay Lines

Note: Each neuron in the array is contacted by axons from each ear. See page 168 in the Auditory Localization: Physiology entry for additional information (pp. 167–170).

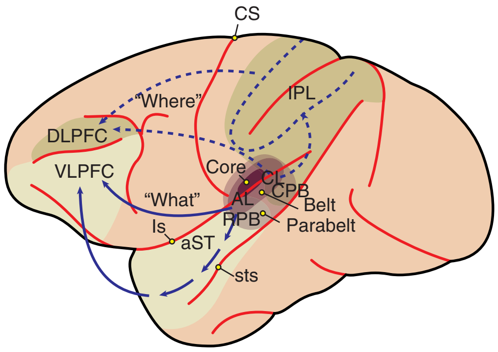


Figure 2 *Auditory Processing: Central*—Schematic Illustration of Hierarchical Processing in the Auditory Cortex of the Rhesus Monkey and Processing Streams for "What" and "Where"

Notes: Auditory signals arrive first in the core areas (consisting of the primary auditory cortex, A1, and two rostral areas) from subcortical inputs in thalamus and brainstem. Activity then propagates to the belt (AL, CL, among others) and parabelt areas (rostral and caudal, RPB and CPB, respectively), which give rise to two pathways projecting to two largely segregated regions in the prefrontal cortex (PFC): the ventrolateral and dorsolateral (VLPFC and DLPFC) regions, respectively. The ventral processing stream is also relayed through the anterior superior temporal (aST) cortex, where regions or patches specialized for the processing of voices and communication sounds have been found in both monkeys and humans. The dorsal stream, in addition to its projection to the DLPFC, is relayed also through the inferior parietal lobule (IPL) of the posterior parietal cortex (PPC, particularly the ventral intraparietal region, VIP). Rich back-projections exist from the prefrontal cortex to the PPC as well as to the aST (not shown). See the Auditory Processing: Central entry for additional information (pp. 176–180).

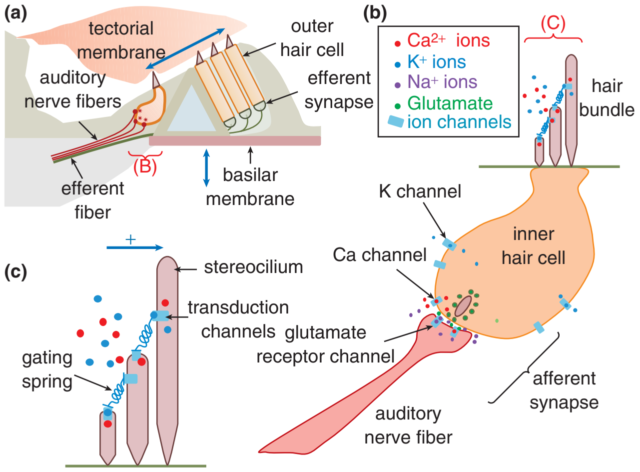


Figure 3 *Auditory Receptors and Transduction*—Hair Cells of the Mammalian Cochlea

Notes: (a) Cross-section through the sensory epithelium (organ of Corti) in the middle chamber of the cochlea. (*Blue arrows*) Sound moves the basilar membrane up and down, bending the hair bundles of inner and outer hair cells against the tectorial membrane. (b and c) Ion channels involved in the afferent flow of electrical signals from the ear to the brain: (b) is a higher-magnification view of the inner hair cell in (a), and (c) is a higher-magnification view of the hair bundle in (b). Deflection of the hair bundle toward its tall edge (+, *blue arrow*) opens mechano-electrical transduction channels, through which potassium (K^+) and calcium (Ca^{2+}) ions enter the stereocilium and make the hair cell more positive. The voltage change activates voltage-gated channels that are selective for K^+ or Ca^{2+} ions. Entry of Ca^{2+} through Ca channels activates release of glutamate by the hair cell onto the afferent nerve ending. Binding of glutamate to glutamate-receptor channels opens the channels, producing an influx of positive ions, which triggers action potentials. See the Auditory Receptors and Transduction entry for additional information (pp. 183–186).

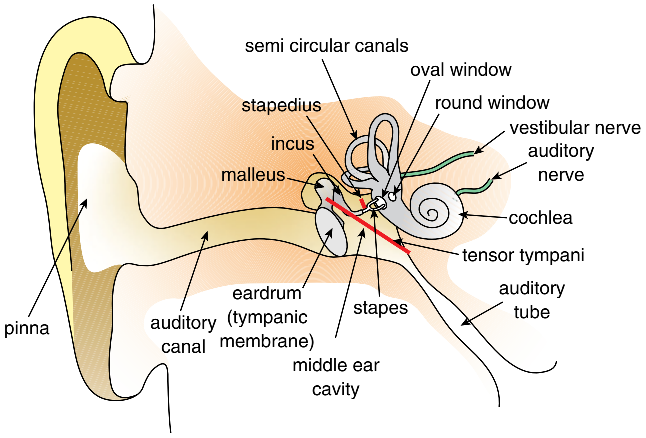


Figure 4 *Auditory System: Structure* (1)—Structure of the Ear

Source: Modified from a Wikipedia illustration by Dan Pickard, with permission.

Notes: Muscles are shown in red and nerves in green. See the Auditory System: Structure entry for additional information (pp. 194–197).

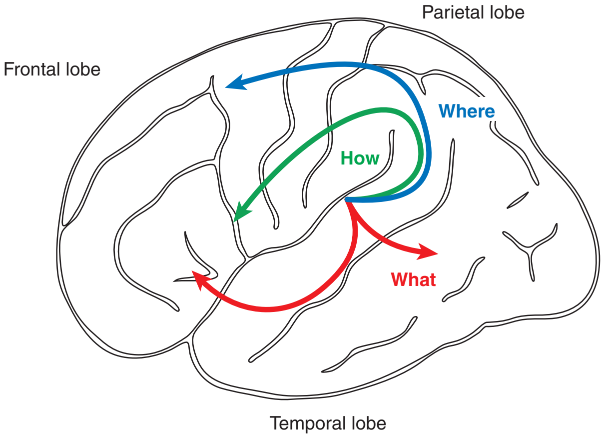


Figure 6 *Speech Perception: Physiological*—Auditory Processing Streams Important for Speech Perception

Source: Adapted from Scott, S. K. (2005). Auditory processing—speech, space and auditory objects. *Current Opinion in Neurobiology, 15,* 197–201.

Note: See the Speech Perception: Physiological entry for additional information (pp. 923–926).

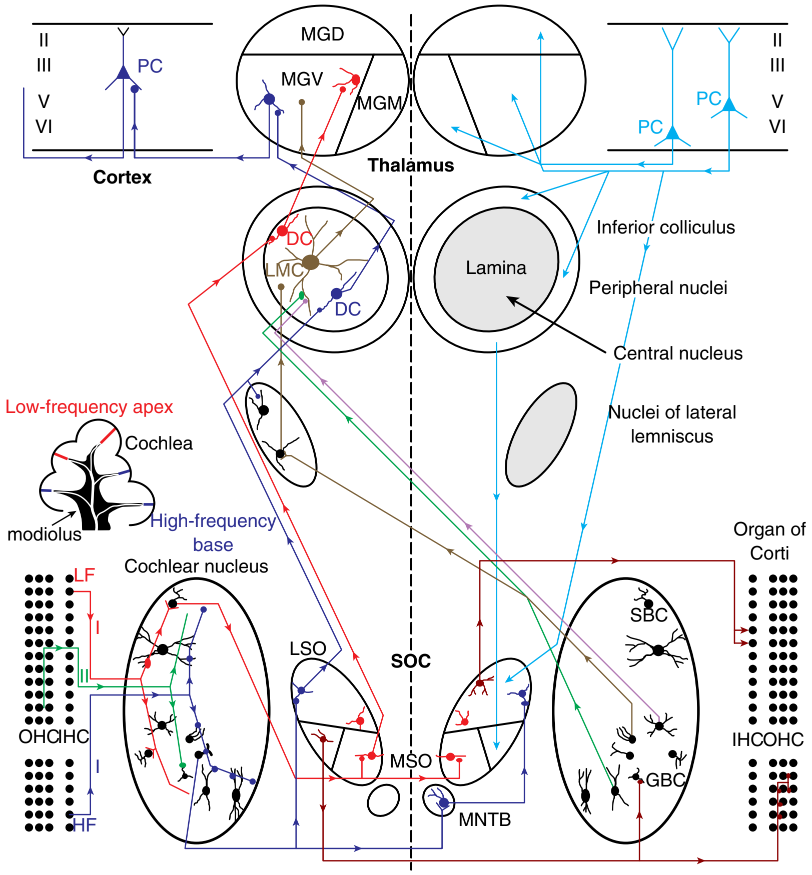


Figure 5 *Auditory System: Structure* (2)—Cell Types and Major Connections Within the Nervous System Processing Auditory Input

Notes: Ascending pathways are mainly shown on the left and descending pathways are mainly shown on the right. The dashed line represents the midline. See the Auditory System: Structure entry for additional information (pp. 194–197).

Cutaneous

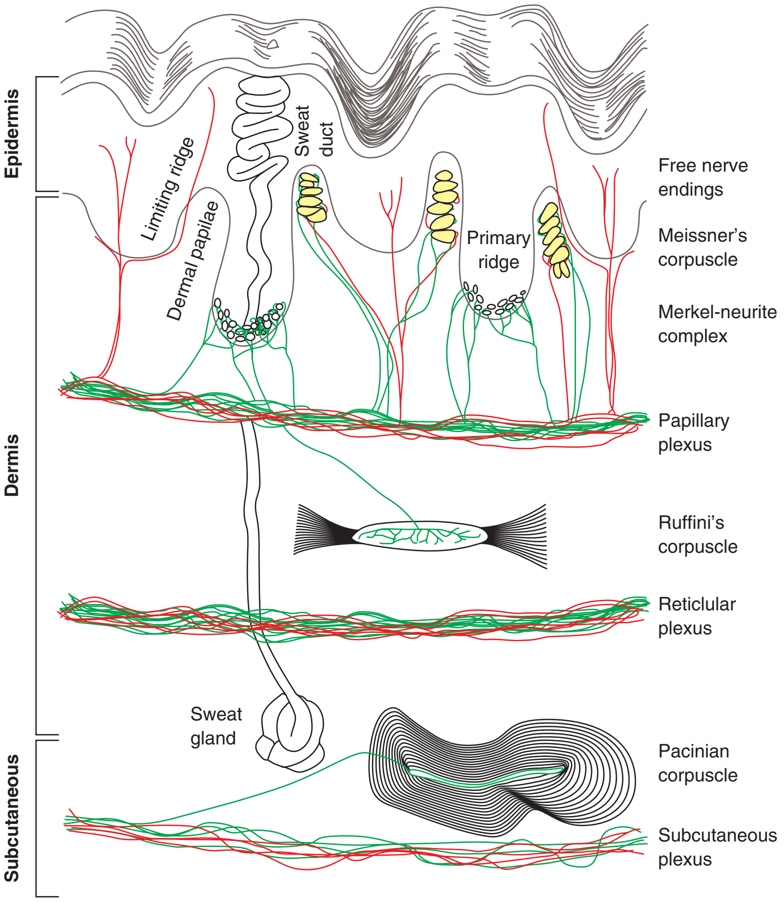


Figure 7 *Cutaneous Perception: Physiology*—Cross-Section Diagram Showing the Different Kinds of Afferent Receptors Innervating the Glabrous Skin of the Primate

Source: Adapted from Johnson, K. O. P. (2002). Neural basis of haptic perception. In H. Pashler & S. Yantis (Eds.), *Steven's handbook of experimental psychology: Vol. 1. Sensation and perception* (3rd ed., pp. 537–583). New York: Wiley.

Notes: The receptor endings that are associated with cutaneous processing are the Meissner's corpuscle, which is the receptor ending for the rapidly adapting afferents (RA); the Merkel-neurite complex, which is the endings for the slowly adapting type 1 afferents (SA1); the Pacinian corpuscle, which is the ending for the Pacinian afferents (PC); and the Ruffini's corpuscle, which was once thought to be the ending for the slowly adapting type 2 afferents (SA2). The free-nerve endings provide the inputs for the pain, temperature, and itch afferents. See the Cutaneous Perception: Physiology entry for additional information (pp. 348–353).

Olfaction

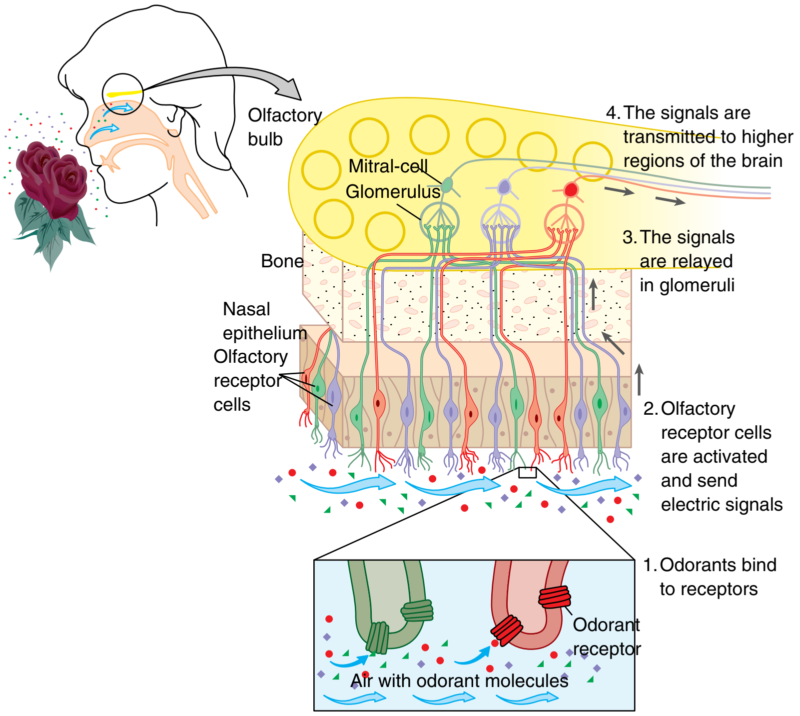


Figure 8 *Olfaction*–Olfactory Epithelium and Bulb

Notes: Schematic of the olfactory epithelium showing how the axons of the olfactory receptors cells that express the same receptors project to common glomeruli within the olfactory bulb. See the Olfaction entry for additional information (pp. 657–661).

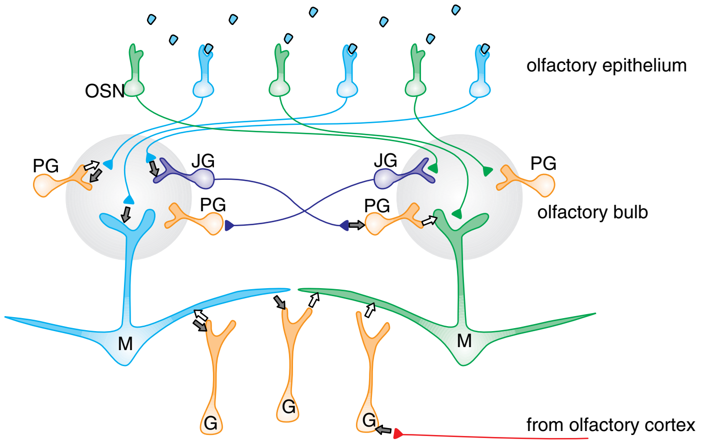


Figure 9 *Olfaction: Feature Detection and Integration*—Basic Circuit Outlining the Flow of Excitation and Inhibition in the Olfactory Bulb

Notes: When excited by odorants, OSNs activate a network of intrinsic and output neurons associated with a single glomerulus. This circuit makes use of both feedback and lateral inhibition to enhance the contrast between similarly responding glomeruli. Abbreviations: OSN = olfactory sensory neuron; PG = periglomerular cell; JG = juxtaglomerular cell; M = mitral/tufted cell. Black arrows indicate excitatory synapses and white arrows indicated inhibitory synapse. See the Olfaction: Feature Detection and Integration entry for additional information (pp. 668–671).

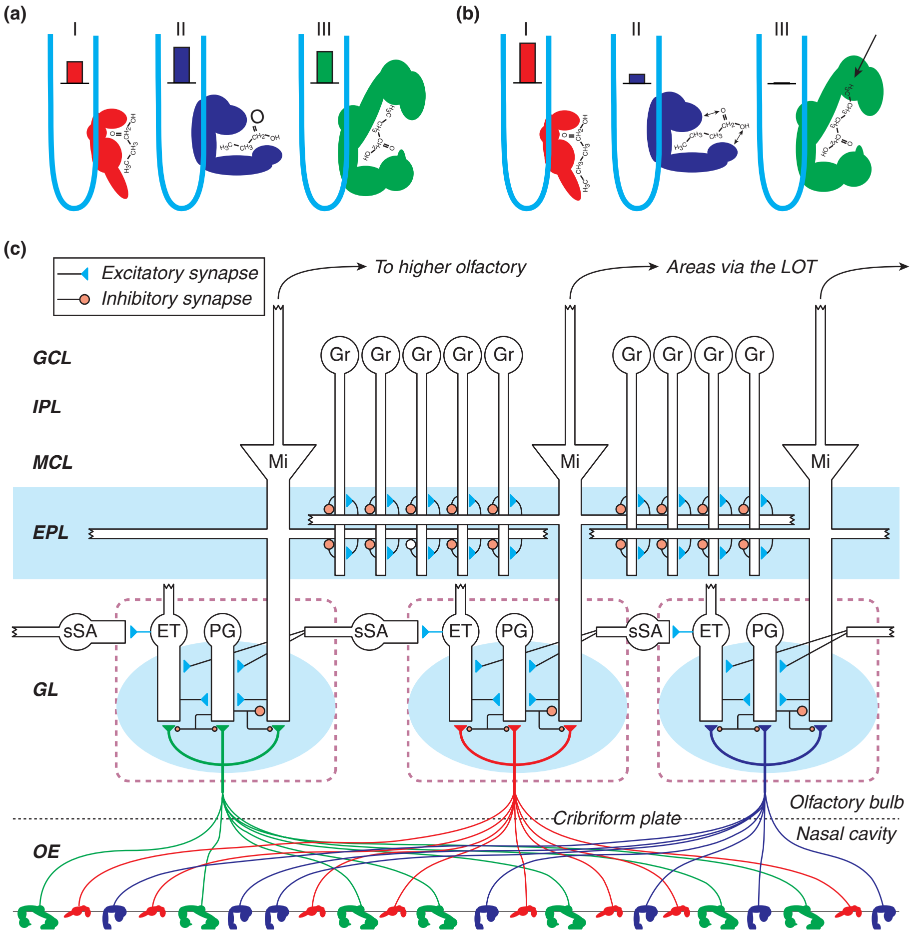


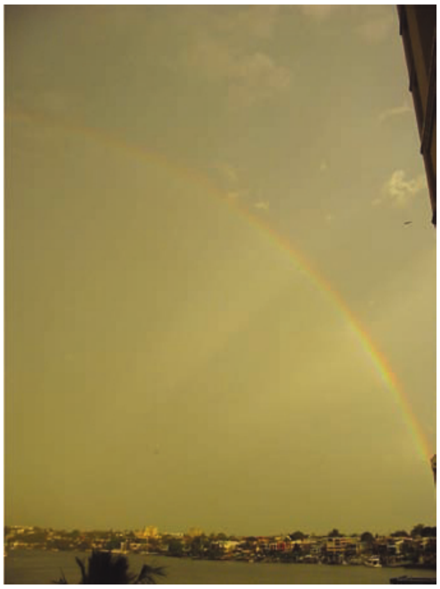
Figure 10 *Olfactory Bulb: Functional Architecture*—Coding and Anatomy of the Olfactory System

Note: See pages 680–681 in the entry Olfactory Bulb: Functional Architecture for a complete description.

Vision

(a)

(b)



(c)

Figure 11 *Atmospheric Phenomena*–Some Atmospheric Phenomena

Sources: Figure 1(a) was kindly supplied by Dr. Urte Roeber, who retains its copyright. Figure 1(b) © Photograph by "Σ64." Retrieved January 16, 2009, from http://commons.wikimedia.org/wiki/File:Brocken-tanzawa.JPG. This file is licensed under the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation (http://commons.wikimedia.org/wiki/Commons:GNU_Free_Documentation_License); with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

Notes: (a) Aerial perspective makes distant objects appear to be lighter and to have less contrast than near objects, here seen in the two slopes, both of which are covered in essentially identical vegetation and are identically illuminated. (b) Four antisolar phenomena: The shadow of the photographer is a *spectre*. It is surrounded by a bright *halo* from back-scattered sunlight. The halo is surrounded by a *glory*, the colored rings. The photographer's shadow also shows *anticrepuscular* rays. (c) Two antisolar phenomena: The fan shape of shadow rays are anticrepuscular rays. These arise from parallel light from the sun being blocked by clouds in the west, creating parallel shadows that receding from the viewer in the east. The rainbow arises from refraction and total internal reflection of sunlight in raindrops. This display of anticrepuscular rays and rainbow is called *rainbow spokes*. See the Atmospheric Phenomena entry for additional information (pp. 67–71).

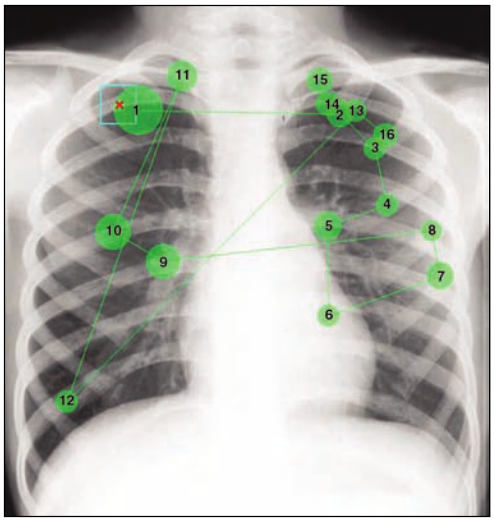


Figure 12 *Attention and Medical Diagnosis*—Expert's Scanpath

Notes: A lesion has been identified by a mouse click (red cross in blue box) at fixation 1 in the image. The first fixation is at the location of the nodule, which suggests the lesion must have been identified in the global look. The size of the green circle is proportional to the length of time of each fixation, with the first fixation being 1.8 seconds. See the Attention and Medical Diagnosis entry for additional information (pp. 119–121).

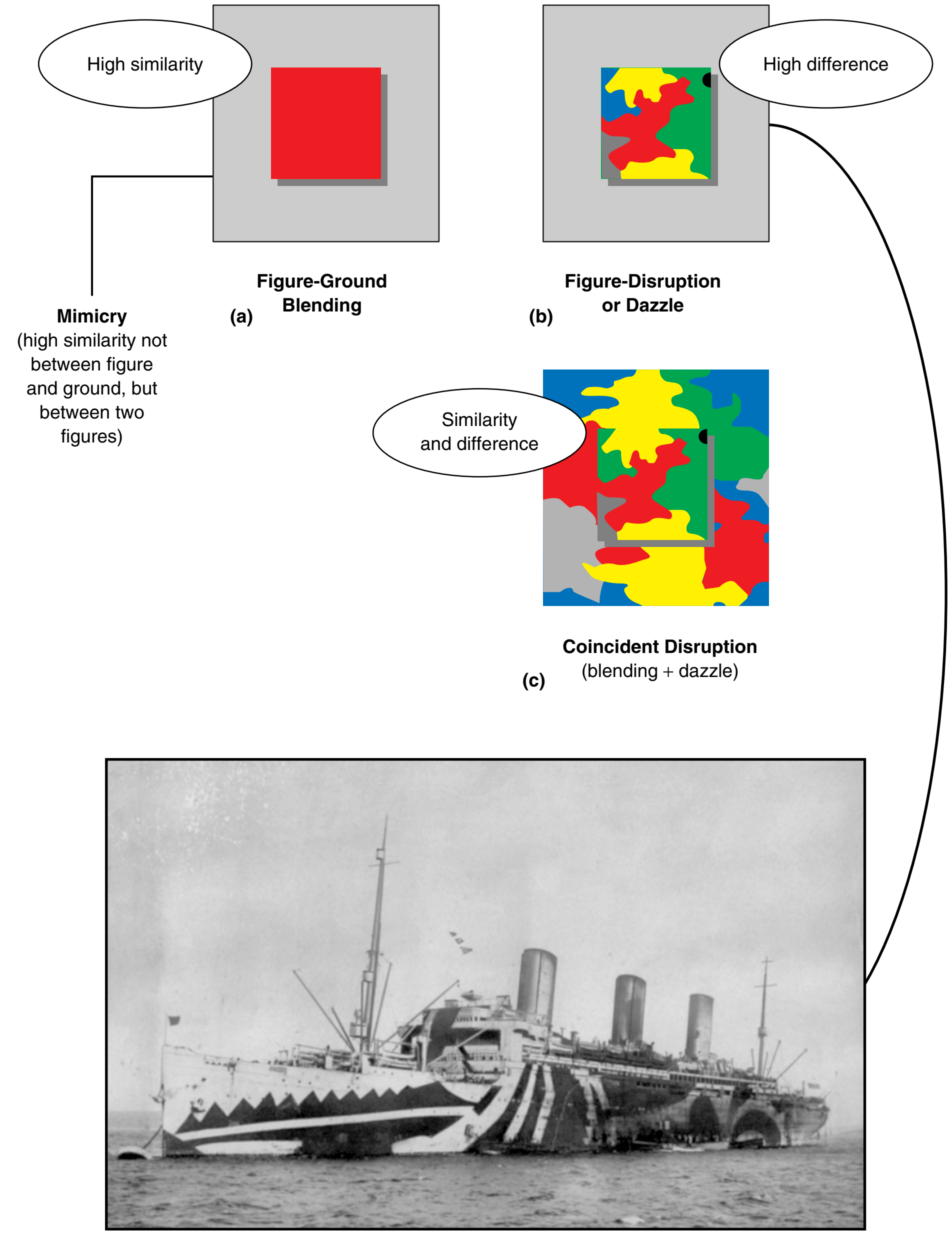
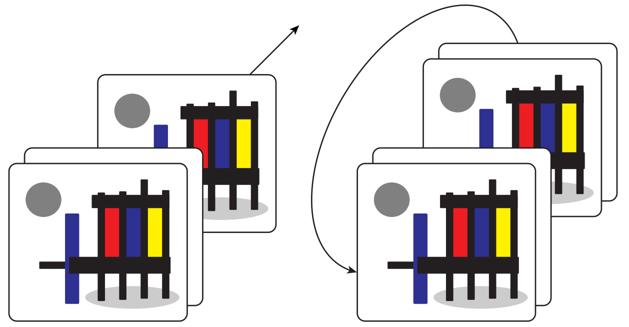


Figure 13 *Camouflage*—Three Major Categories of Camouflage

Source: Illustration © by Roy R. Behrens, 2008. Reprinted with permission.

Notes: Although there are any number of ways to categorize examples of camouflage, perhaps the three most common kinds are figure-ground blending (a), in which an object blends in with its setting (also sometimes known as crypsis); figure disruption (b and photo below it), in which the surface of the figure is broken up by highly contrasting components, a method that was used widely for ship camouflage during World Wars I and II; and coincident disruption (c), in which the figure is broken up, but parts of it blend in with the background. In the system shown here, mimicry is said to be a subcategory of blending or high similarity camouflage. See the Camouflage entry for additional information (pp. 233–236).



(a) One-shot paradigm **(b)** Flicker paradigm

Figure 14 *Change Detection*—Example of Method Used to Study Change Detection

Notes: Here, a gap-contingent technique makes the change at the same time a brief blank appears. Two ways of measuring performance are illustrated. (a) One-shot paradigm. The observer views a single alternation of the stimuli; performance is measured by the accuracy of detection (or identification) of the change. (b) Flicker paradigm. The observer views a continual cycling of stimuli; performance is measured by the time taken until the change is detected (or identified). Both measurement paradigms can also be applied to other techniques, such as changes made during an eye movement or a shift in the image. (The change between the front and back images is in the width of the horizontal black bar under the colored panels.) See the Change Detection entry for additional information (pp. 241–244).

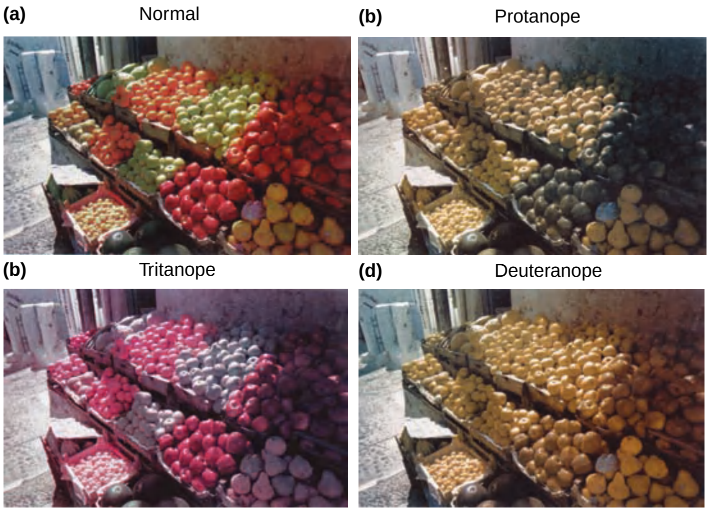
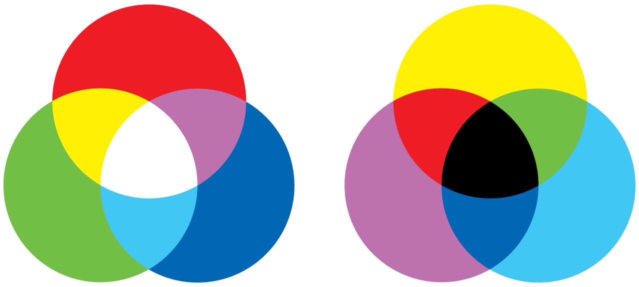


Figure 15 *Color Deficiency*—Perceptual Consequences of Color Vision Defects

Source: Gegenfurtner, K. R., & Sharpe, L. T. (1999). *Color vision: From genes to perception*. New York: Cambridge University Press. Reproduced with permission of Cambridge University Press.

Notes: Here is a simulation of how a scene from a fruit market is perceived by (a) a normal trichromat, (b) a protanope, (c) a deuteranope, and (d) a tritanope. Each color vision deficiency shows greatly reduced chromatic discrimination compared with that of a normal trichromat. See the Color Deficiency entry for additional information (pp. 257–261).



Additive color mixing Subtractive color mixing

Figure 16 *Color Mixing*—Additive and Subtractive Color Mixing (1)

Notes: The left panel simulates the additive combination of red, green, and blue lights. The right panel simulates the subtractive combination of yellow, purple, and cyan pigments. See the Color Mixing entry for additional information (pp. 262–264).

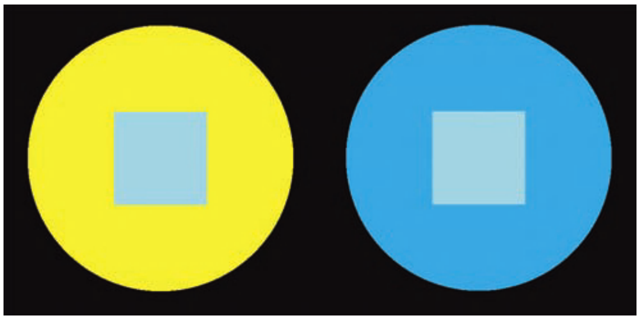


Figure 17 *Color Mixing*—Color Appearance (2)

Notes: The smaller squares are physically the same (i.e., they are printed with the same ink), but their colors appear different. The differences arise because of the surrounding colors, which induce color changes in the appearance of the central squares. See the Color Mixing entry for additional information (pp. 262–264).

Figure 18 *Color Perception*—Image Decomposed Into Its Luminance or Chromatic Components (1)

Notes: The grayscale image (left) retains much of the fine spatial detail, depth, and shading information in the scene. The chromatic images (center and right) captures the material differences between the objects much better and more clearly delineates the objects in terms of these differences and despite variations in shadows and shading. Thus, adding color helps viewers perceive both what the objects are and where they are. See the Color Perception entry for additional information (pp. 266–270).

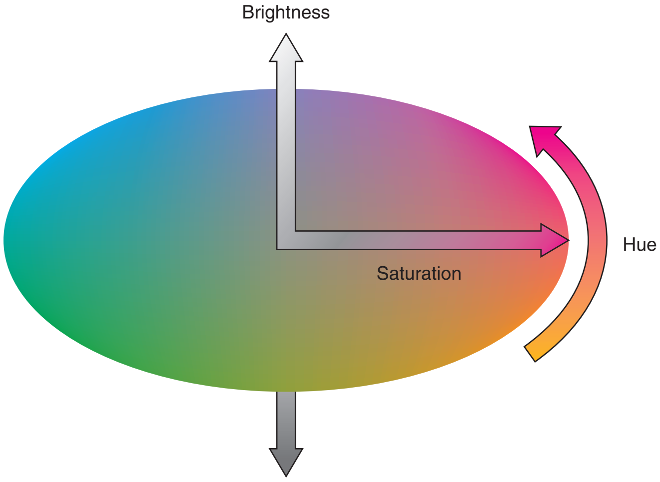


Figure 19 *Color Perception*—The Appearance of an Unrelated Color Can Be Described by Its Brightness, Hue, and Saturation (2)

Notes: These three attributes are arranged in perceptual color spaces to represent colors according to how they differ from a neutral gray. See the Color Perception entry for additional information (pp. 266–270).

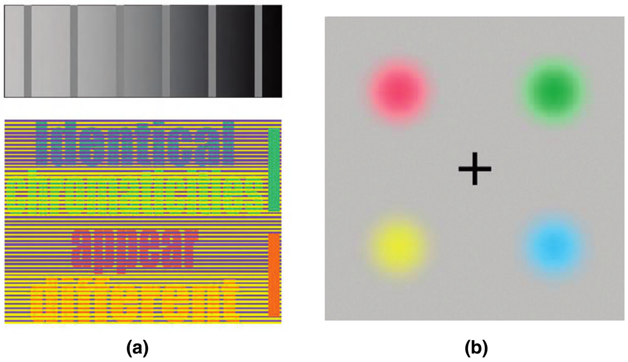


Figure 20 *Color Perception*—Examples of Contextual Effects in Color Appearance (3)

Notes: (a) The thin bars in the upper image are all the same printed gray but appear darker or brighter because of the contrast difference with the neighboring background. Similarly, in the lower figure, the top two and bottom two words are the same physical chromaticity (green and orange, as shown by the bars on the right), yet they appear dramatically different when they are interleaved with either the purple or yellow lines (image courtesy of Patrick Monnier, Department of Psychology, Colorado State University). In (b), as you stare directly at the cross in the center, the colored patches will fade away and may disappear entirely. This occurs because each area of the retina adapts over time to the color falling on it, a process that keeps color appearance centered around the average stimulus we are exposed to. See the Color Perception entry for additional information (pp. 266–270).

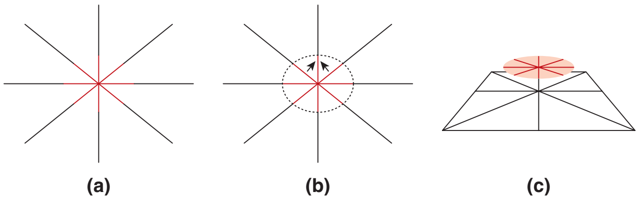


Figure 21 *Computational Approaches*—Neon Color Spreading

Notes: (a) An image that shows the perceptual effect of neon-color spreading. (b) Construction of the invisible fence (dotted line) based on cue discontinuity and border linking. Diffusion spreads the red hue over the surface (arrows) but is contained by the invisible fence. (c) Presumed mental interpretation of a tinted red surface on top of a surface with black lines. See the Computational Approaches entry for additional information (pp. 278–283).

Figure 22 *Digital Imaging*—Color Filter Array (CFA) Overlaid Onto a Charge Coupled Device (CCD) Sensor

Notes: Each CCD element records a limited range of wavelengths, corresponding to either red (gray), green (light gray), or blue (dark gray). A full three-channel RGB color image is created by interpolating the missing color pixels by, for example, averaging the recorded values. See the Digital Imaging entry for additional information (pp. 362–366).

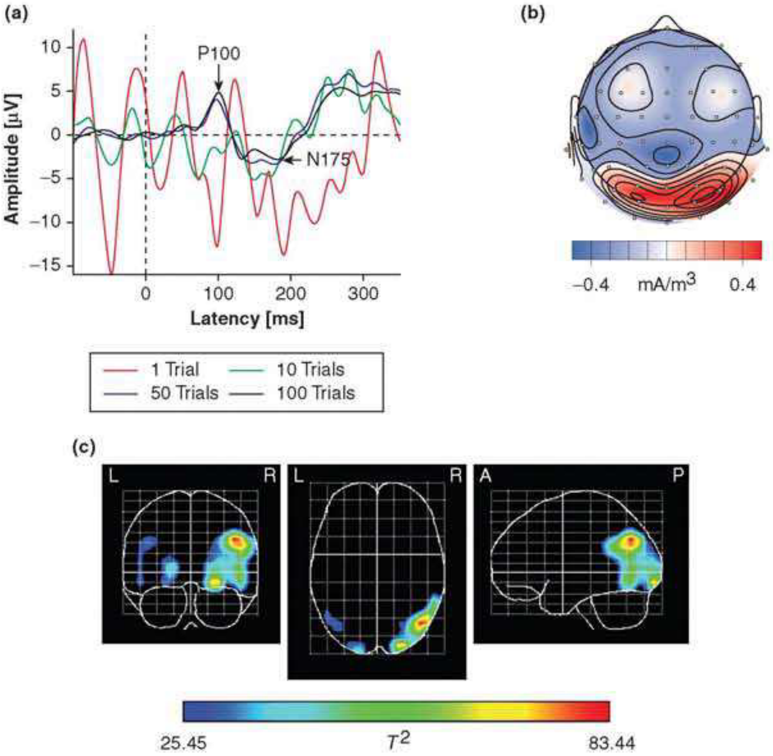


Figure 23 *Evoked Potential: Vision*—Visual Evoked Potential Records

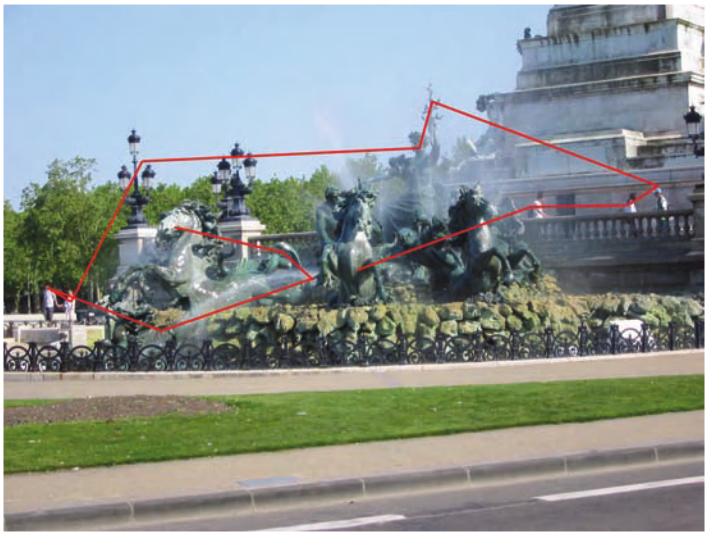
Source: This figure is based on data from Roeber, U., Widmann, A., Trujillo-Barreto, N. J., Herrmann, C. S., O'Shea, R. P., & Schröger, E. (2008). Early correlates of visual awareness in the human brain: Time and place from event-related brain potentials. *Journal of Vision, 8*(3), 1–12.

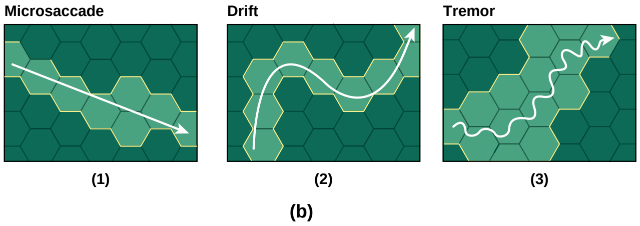
Notes: (a) A person's EEG activity averaged from three electrodes over the right parieto-occipital scalp as a grating presented to one eye changed its orientation (at time 0) from being perpendicular to an identical grating presented to the other eye to being the same as the grating in the other eye. The VEP requires averaging of many EEG records from many trials showing the same visual event. After one trial (the red trace), the EEG activity is essentially random, the VEP being obscured by other activity in the brain. After 10 trials (green trace), the average EEG activity is still essentially random, although its variance is much reduced as the random activity on one trial tends to cancel out the random activity on another. After 50 trials (blue trace), the average EEG activity is quite regular, and the VEP is clearly evident, the random activity having been largely averaged out. After 100 trials (black trace), the VEP is clearly revealed. (b) The scalp current densities computed over 61 scalp electrodes from 12 participants in the same experiment. Changes evoke larger currents over the right occipital area. (c) The sources of the electrical activity computed by VARETA (variable resolution electromagnetic tomography), shown in a "glass brain." It occurs mainly in the right ventrolateral occipital area. See the Evoked Potential: Vision entry for additional information (pp. 399–401).

Figure 24 *Eye Movements and Action in Everyday Life*—The RIT Wearable Eyetracker, Showing the Scene Camera and Eye Camera Mounted on a Lightweight Glasses Frame

Source: M. Hayhoe and D. Ballard, 2005.

Notes: (a) The backpack carries a battery and video camera, which tracks eye position relative to the scene the person is observing. (b) Composite image of a scene from the head-mounted video camera while the subject makes a peanut butter and jelly sandwich. The images from the camera were integrated over different head positions. The fixations are indicated by the yellow circles in the color insert. Circle diameters reflect the duration of the fixation. See the Eye Movements and Action in Everyday Life entry for additional information (pp. 430–433).





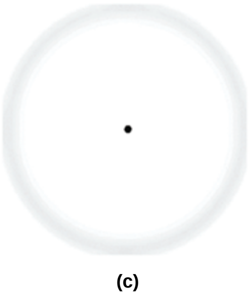


Figure 25 Eye Movements During Fixation

*Sources: (*a) Henderson, J. M. (2009). Used with permission. (b) E. Bruce Goldstein, 2009. (c) Modified from Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience, 5,* 229–240.

Notes: (a) Eye movements to a picture of a fountain. Each line is a saccadic eye movement. Each saccadic eye movement ends in a fixation, where the eye rests briefly before the next saccadic eye movement. (b) Three types of fixational eye movements: movements of the eye that occur during fixation. (c) Troxler fading. In 1804, Swiss philosopher Ignaz Paul Vital Troxler discovered that deliberately fixating on something causes surrounding stationary images to fade away. To elicit this experience, stare at the central dot while paying attention to the surrounding pale ring. The ring soon vanishes, and the central dot appears set against a white background. Move your eyes, and it pops back into view. See the Eye Movements During Fixation entry for additional information (pp. 438–439).

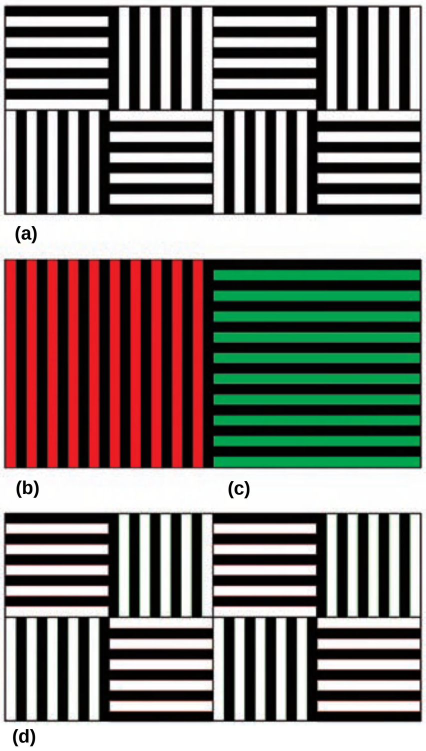


Figure 26 *McCollough Effect*—McCollough Effect Stimuli

Notes: To see the effects illustrated in this figure best, view it from seven times the height of panel (a). (a) The test figure for the McCollough effect. Vertical and horizontal lines look, and are, black-and-white. (b, c) Induction stimuli. Look at the center of (b) for about 10 seconds, then the center of (c) for about 10 seconds; continue looking between (b) and (c) for about five minutes. When you look back at (a) you might see something similar to (d): all of the lines, including the white space between them, appear tinted with faint complementary colors. This stimulus is the same as (a), largely black and white, except that very fine lines of red and green have been added to the edges. These fine lines can be seen by looking closely. See the McCollough Effect entry for additional information (pp. 547–549).

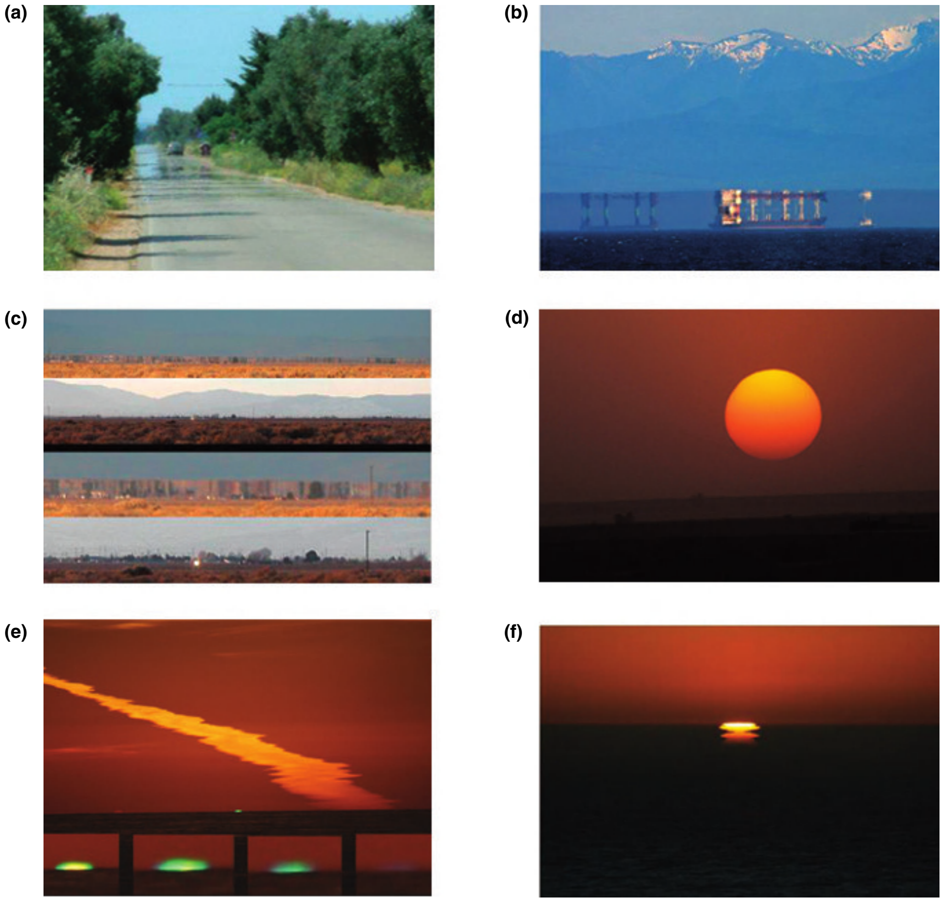


Figure 27 Mirages and Other Atmospheric Phenomena From Refraction

Sources: (a) Photograph by Anton. Retrieved December 19, 2008, from http://upload.wikimedia.org/wikipedia/commons/2/27/Miragerp.jpg. This file is licensed under the Creative Commons Attribution ShareAlike 2.5 License (http://creativecommons.org/licenses/by-sa/2.5/). (b) Photograph by Craig Clements. Retrieved December 19, 2008, from http://commons.wikimedia.org/wiki/File:Superior_mirage_of_the_boats_at_entrance_of_the_harbor_at_Victoria,_British_Columbia,_Canada.jpg. This file is licensed under the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation (http://commons.wikimedia.org/wiki/Commons:GNU_Free_Documentation_License); with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. (c) Image by Alan K. Radecki. Retrieved December 20, 2008, from http://commons.wikimedia.org/wiki/File:Mirage3-12.jpg. This file is licensed under the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation (http://commons.wikimedia.org/wiki/Commons:GNU_Free_Documentation_License); with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. (d) Image taken from a photograph by SSgt Shane Cuomo, U.S. Air Force. Retrieved December 27, 2008, from http://www.dodmedia.osd.mil/Assets/Still/2005/Air_Force/DF-SD-05-04367.jpeg.This image or file is a work of a U.S. Air Force Airman or employee, taken or made during the course of the person's official duties. As a work of the U.S. federal government, the image or file is in the public domain. (e) Image by Mila Zinkova (http://home.comcast.net/~milazinkova/Fogshadow.html). Retrieved December 19, 2008, from http://commons.wikimedia.org/wiki/File:Inferior_Mirage_green_flash.jpg. This file is licensed under the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation (http://commons.wikimedia.org/wiki/Commons:GNU_Free_Documentation_License); with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. (f) Photograph by Mila Zinkova (http://home.comcast.net/~milazinkova/Fogshadow.html). Retrieved December 19, 2008, from http://commons.wikimedia.org/wiki/File:Mock_mirage_of_the_setting_sun_02-08-08-1.jpg. This file is licensed under the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation (http://commons.wikimedia.org/wiki/Commons:GNU_Free_Documentation_License); with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

Notes: (a) An inferior mirage on a road: there is an (inverted) image of the blue sky on the road and a distorted, inverted image of the trees. (b) A superior mirage: The superior mirage is the inverted image above the erect image of the ship. Above that is another, narrow, erect image of the ship, also a superior mirage. (c) A fata morgana: This mirage, shown in four panels, involves numerous alternately inverted and erect superior mirages. In the top panel is the fata morgana. In the next panel is the same scene without any mirage. The next two panels show magnified versions of the first two. (d) The flattened sun: As the sun sets, it becomes flattened vertically. (e) The green flash: The upper panel shows an overview of the sun almost completely set showing only a green part. The small lower panels show a magnified view of the sun in sequence as it disappeared below the horizon. (The diagonal cloud is from a contrail.) (f) A mock mirage giving a paper-lantern sun. See the Mirages entry for additional information. (pp. 555–557).

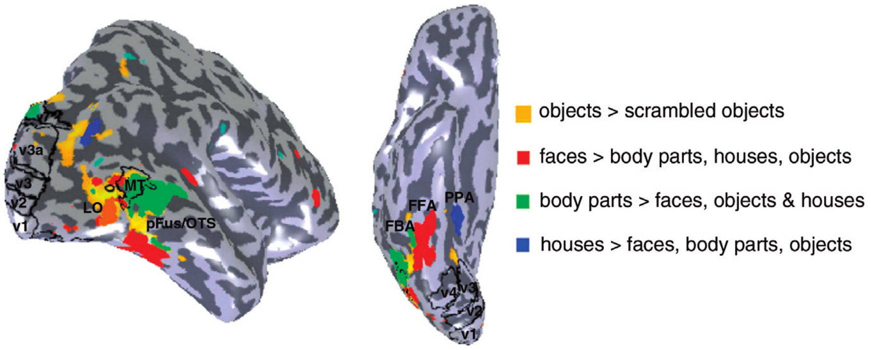


Figure 28 *Object Perception: Physiology*—Functional Magnetic Resonance Imaging (fMRI) Responses to Various Stimuli

Notes: Object (yellow), face (red), body part (green), and house (blue) selective regions presented on an inflated cortical surface of a right hemisphere of a representative subject. Dark grays indicate sulci and lighter grays indicate gyrii. *Left:* lateral view. *Right:* ventral view. Orange indicates regions that are both face and object-selective. Abbreviations: LO = lateral occipital; pFus/OTS = posterior fusiform/occipitotemporal sulcus; MT = mid-temporal motion selective region; FBA = fusiform body part area; FFA = fusiform face area; PPA = parahippocampal place area. Boundaries of early visual areas V1, V2, V3, V3a, and V4 are shown in black. See the Object Perception: Physiology entry for additional information (pp. 648–653).

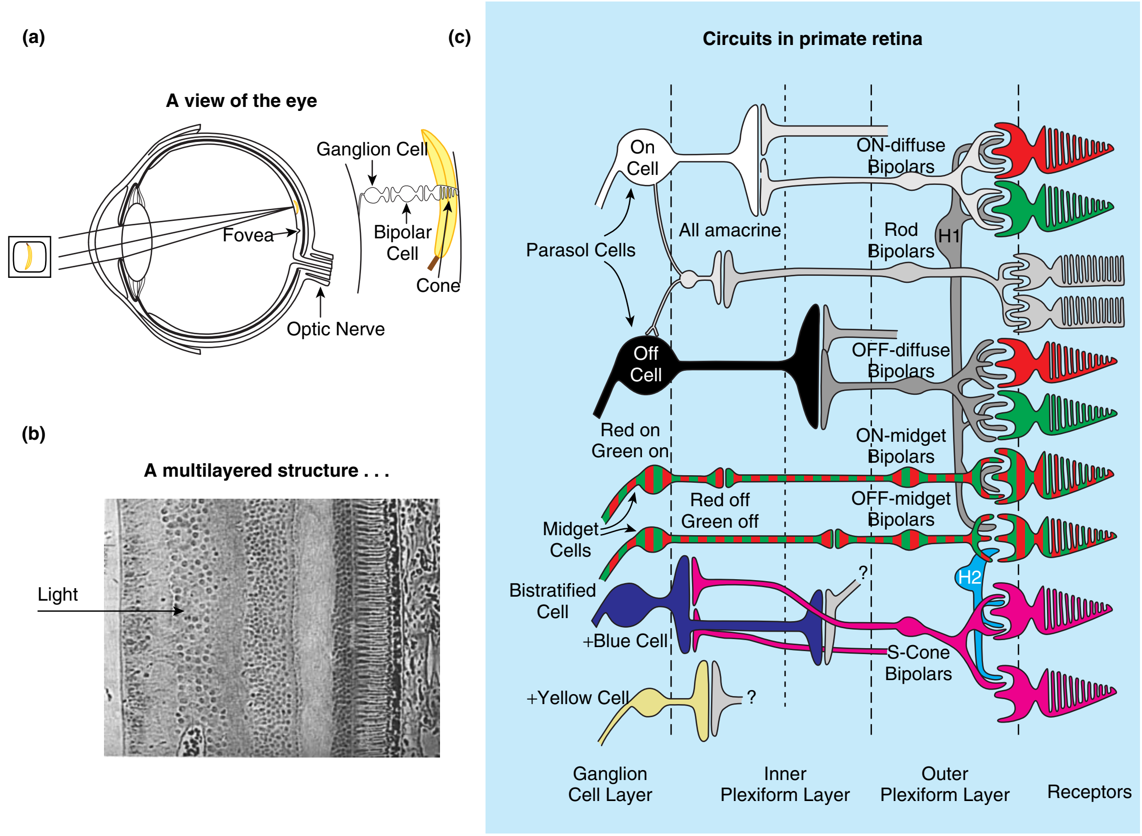


Figure 29 *Retinal Anatomy*—The Eye and Retina

Notes: (a) Sketch of the eye. The retina lies at the back of the eye. The fovea is the region of central vision. The basic retinal circuit (inset) consists of photoreceptors (anchored to the pigment epithelium) that give their output to the bipolar cells that give their output to the ganglion cells, which send their output to the brain. (b) The retina shows a regular structure with different lamina; light must pass through them to reach the receptors. (c) A wiring diagram of the different channels in primate retina. Different channels mediate achromatic, red-green, and blue-yellow signals. Each has the same pattern—cone to bipolar cell to ganglion cells, but there are differences in detail. See the Retinal Anatomy entry for additional information. (pp. 868–871).

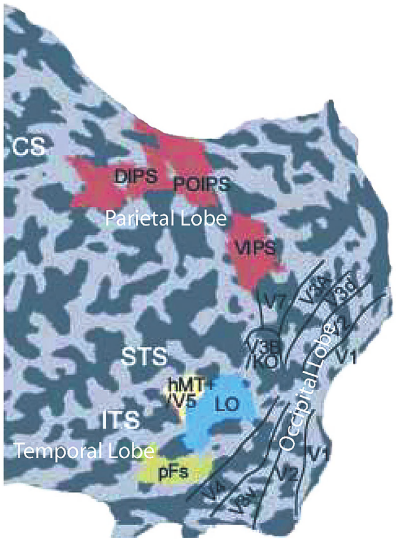
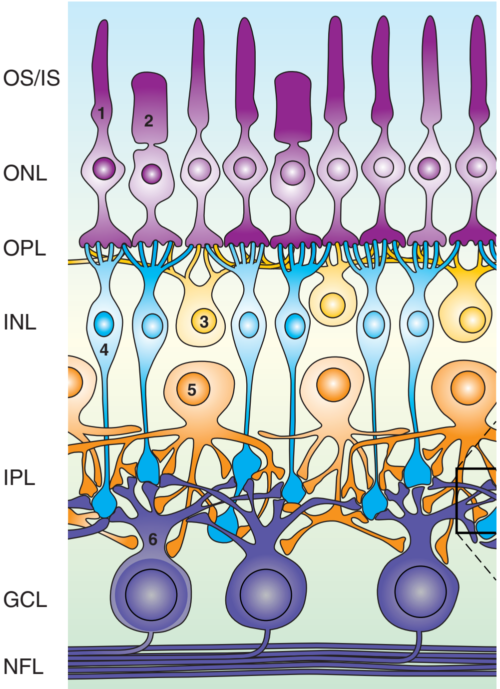


Figure 30 *Visual Processing: Extra-Striate Cortex*—Areas in Extrastriate Cortex

Notes: Activation maps on a flattened cortical surface showing retinotopic areas (V1, V2, V3, V4, V7, V3B/KO), the human motion complex (hMT+/V5), the lateral occipital complex (LOC with subregions: LO = lateral occipital; pFs = posterior fusiform sulcus), and 3-D shape-related areas in the parietal cortex (VIPS = ventral inferior parietal sulcus, POIPS = junction of posterior occipital and inferior parietal sulcus, DIPS = dorsal inferior parietal sulcus). Left hemisphere is shown. Sulci are coded in darker grey than the gyri. Major sulci are labeled as follows: STS = superior temporal sulcus; ITS = inferior temporal sulcus; CS = central sulcus. See the Visual Processing: Extra-Striate Cortex entry for additional information. (pp. 1096–1099).



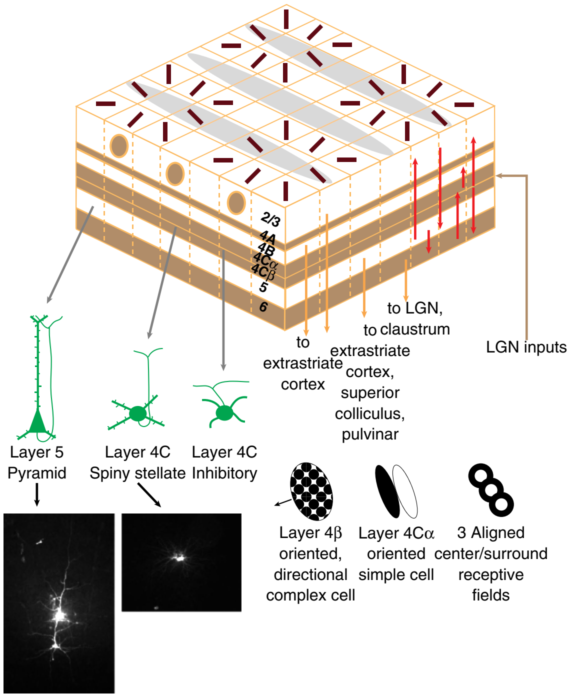


Figure 31 *Visual Processing: Primary Visual Cortex*—Functional Organization of the Primary Visual Cortex, V1

Notes: V1 is a three-dimensional structure containing six main layers. Within these layers, three major cell types—pyramidal cells, spiny stellate cells, and inhibitory cells (shown in green; example intracellular fills of a pyramidal and stellate neuron are illustrated below each)—provide a specific organization of intrinsic connections (red arrows) and extrinsic connections (orange arrows). The organization of lateral geniculate nucleus (LGN) inputs and intrinsic connections underlie the construction of cortical receptive fields (bottom right illustrates a progression from aligned center/surround receptive fields to an oriented simple cell receptive field and a direction selective complex cell receptive field). Dashed orange lines illustrate the boundaries of cortical columns, which contain neurons with similar response properties (e.g., all neurons in a column are tuned to the same stimulus orientation, illustrated here by oriented black bars drawn on the surface). Across the cortex, columns are organized in a precise fashion to form maps of visual space (retinotopy), ocular dominance (illustrated by grey stripes on the cortical surface), and orientation preference (illustrated by oriented bars). See the Visual Processing: Primary Visual Cortex entry for additional information. (pp. 1099–1101).

Figure 32 *Visual Processing: Retinal*—Neuroanatomy of the Retina

Source: Wassle, H. (2004). Parallel processing in the mammalian retina. *Neuroscience Nature Reviews, 5,* 747–757.

Notes: Cell types: Cones (#1), rods (#2), horizontal cells (#3), bipolar cells (#4), amacrine cells (#5), ganglion cells (#6). Abbreviations: OS/IS = outer segments and inner segments; ONL = outer nuclear layer; OPL = outer plexiform layer; INL = inner nuclear layer; GCL = ganglion cells; NFL nerve fibers. See the Visual Processing: Retinal entry for additional information. (pp. 1101–1104).

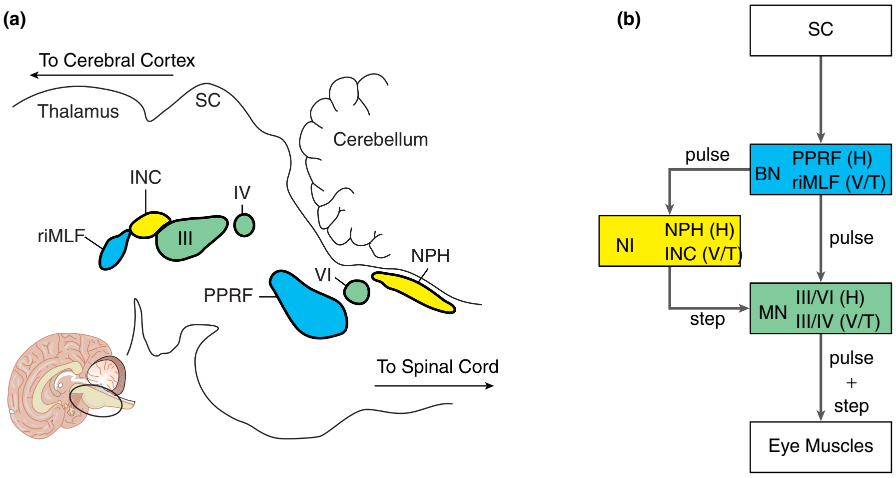


Figure 33 *Visual Processing: Subcortical Mechanisms for Gaze Control*—Brain Stem Pathways for Gaze Control

Sources: (a) V. Henn, J. A. Buttner-Ennever, and K. Hepp, 1982. (b) Author.

Notes: (a) Midsagittal view of the primate brain stem. Burst neurons are shown in blue, neural integrators in yellow, and motoneurons in green. The superior colliculus (SC), thalamus, and cerebellum are also shown. Inset shows the location of the brain stem. (b) A schematic of gaze shift commands traveling through the brain stem. Color scheme same as in (a). Horizontal components are indicated by (H), while vertical/torsional components are indicated by (V/T). SC = superior colliculus; BN = burst neurons; NI = neural integrator; MN = motoneurons; III = oculomotor nucleus; IV = trochlear nucleus; VI = abducens nucleus. See the Visual Processing: Subcortical Mechanisms for Gaze Control entry for additional information (pp. 1104–1107).

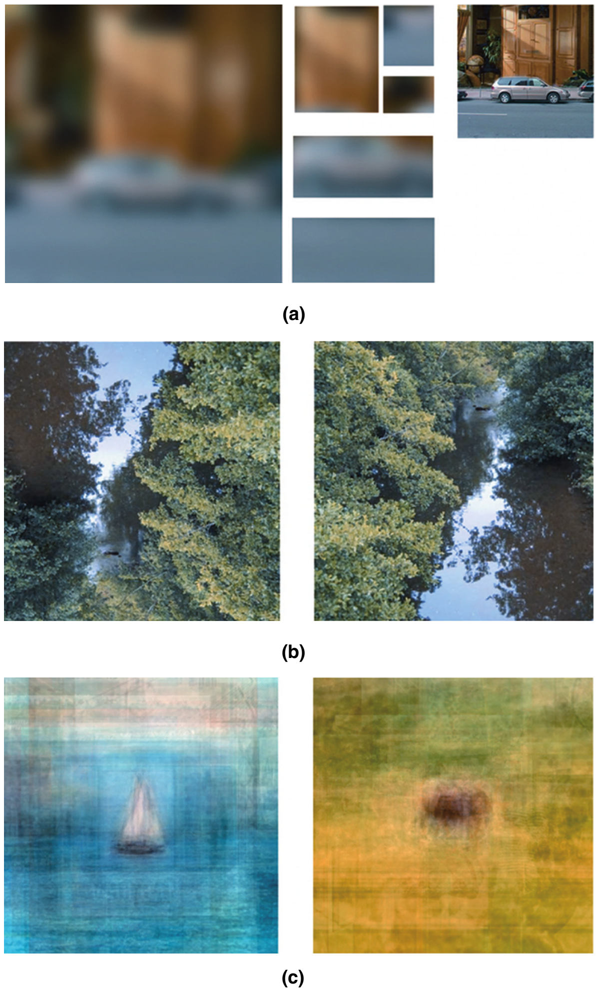


Figure 34 *Visual Scene Perception*—Perceiving Visual Scene Stimuli

Source: Courtesy of Antonio Torralba, MIT.

Notes: (a) The blurry image looks like a street scene, even if the objects cannot be recognized in isolation. (b) A forest scene that changes interpretation with image orientation. (c) Average images centered on pictures of sailboats (left) and cows (right). See the Visual Scene Perception entry for additional information (pp. 1111–1116).

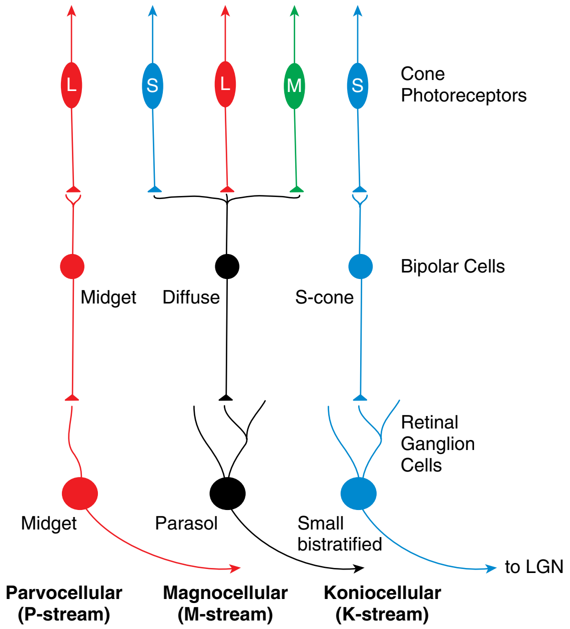


Figure 35 *Visual System Structure* (1)—Foveal Retinal Circuitry

Notes: Three types of cone photoreceptors synapse onto specialized bipolar neurons, which connect to distinct types of retinal ganglion cells depending on their participation in the parvocellular, magnocellular, or koniocellular processing streams. See the Visual System Structure entry for additional information (pp. 1130–1134).

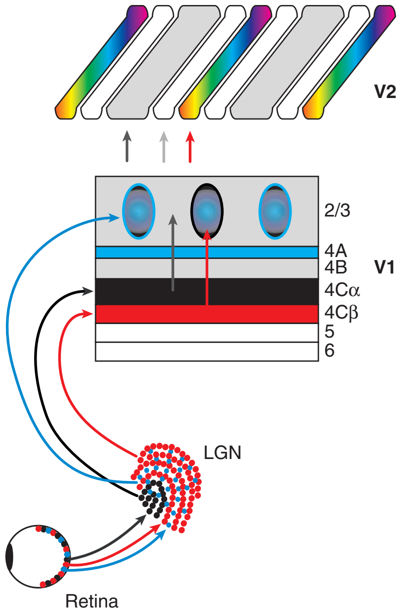


Figure 36 *Visual System Structure* (2)—Parallel Processing Streams of the Early Visual System

Notes: The magnocellular stream illustrated in black, parvocellular stream illustrated in red, koniocellular stream illustrated in blue. Retinal ganglion cells project to lateral geniculate nucleus (LGN) neurons, which project to specific layers within V1. V1 sends motion information to thick stripes in V2 and color/form information to thin and pale stripes in V2. Numbers on the right indicate cortical layers. See the Visual System Structure entry for additional information (pp. 1130–1134).

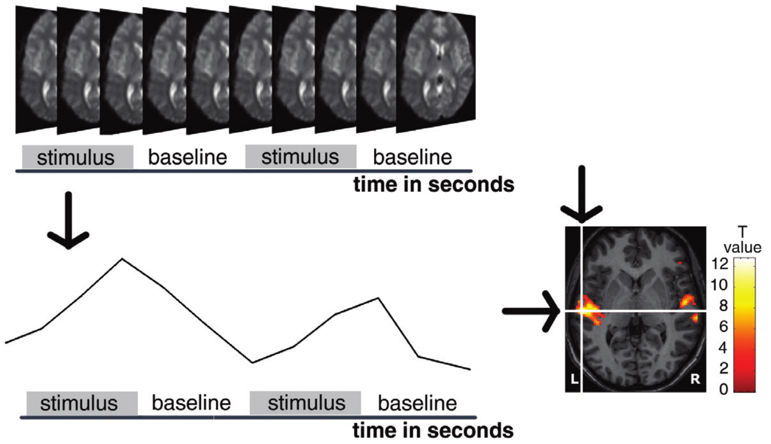


Figure 37 *Brain Imaging*—Several Key Aspects of an (Auditory) fMRI Experimental Design and Image Analysis

Source: Original figure created by D. A. Hall.

Notes: The upper time line represents a segment of the experiment showing how repeated scans are acquired at regular intervals over time, while the participant listens to an alternating sequence of sound stimulation and baseline conditions. Each scan is represented by one horizontal slice through the brain. In this example, each stimulus condition lasts about 9 seconds, with scans occurring every 3 seconds. The lower time line presents a schematic diagram of the stimulus-correlated variability in the MR signal measured in arbitrary units. Such a time course is representative of a voxel (individual point) that responds strongly to the sound stimulus. The spatial distribution of the activated voxels is shown in the right panel, with the most significant voxel highlighted by the crosshairs. The color scale reflects the statistical significance of the activation across the auditory cortex. Note that the activation is overlaid onto the anatomical scan, which has a finer spatial detail than the functional scans shown in the upper time line. In this figure, the activation reflects the greater contralateral auditory cortical response to a sound stimulus that was presented to the right ear (pp. 227–231).

General Perception

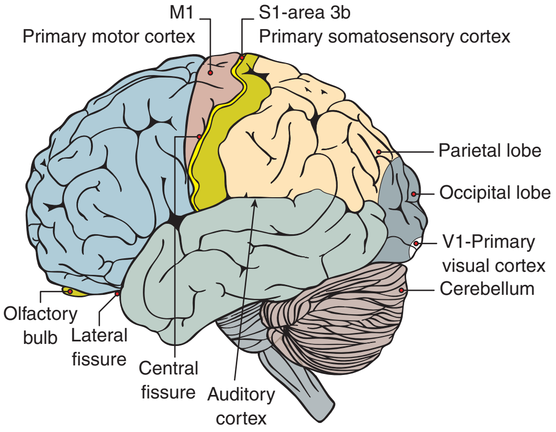


Figure 38 *Cortical Organization* (1)—The Cerebral Cortex of the Human Brain

Notes: The four lobes of the cortex are indicated, as well as parts of primary motor and sensory areas. Because areas extend into the central fissure, only parts of the primary motor cortex (M1 or area 4) and primary somatosensory cortex (area 3b) are visible. The primary auditory cortex is deep in the lateral fissure, and most of the primary visual cortex (V1 or area 17) is hidden in the calcarine fissure on the medial wall of the hemispheres. The human brain has many other cortical areas. See the Cortical Organization entry for additional information (pp. 327–333).

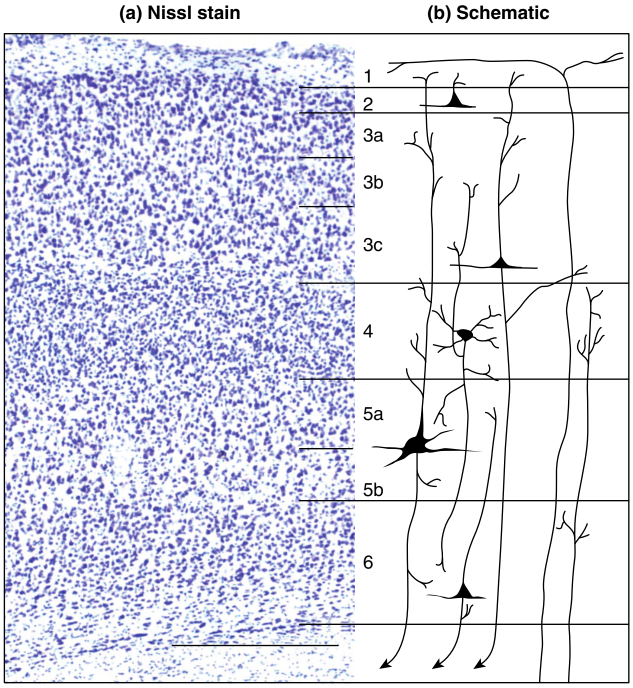


Figure 39 *Cortical Organization* (2)—The Laminar Organization of the Cortex

Notes: (a) A thin section through the thickness of the cortex from the surface above to the fibers (axons that connect the cortex to other regions of the cortex and the brain) below. The Nissl stained section through V1 of a squirrel shows the locations of small, individual neurons as dark dots. Layers (1–6) are distinguished by differences in the arrangements, sizes, and shapes of neurons. (b) A drawing of enlarged neurons (for visibility) and their local connections illustrates how vertical arrays of neurons are densely interconnected. Neurons have large cell bodies, and thin dendrites and axons. Axons entering the cortex and terminating in layers 6, 4, and 1 are shown on the far right side in (b). See the Cortical Organization entry for additional information (pp. 327–333).

A

ABSOLUTE PITCH

Absolute pitch (AP) is the rare ability some people have to instantly and effortlessly name the pitch of musical tones, as readily as most of us name the colors we see. Others are said to possess AP if they can produce (such as by humming, whistling, or singing) a tone given its name. Suppose you are at a concert by your favorite rock group and the guitarist begins to play the opening strains of one of the group's biggest hits. Within a few notes, the entire audience recognizes the song and they jump to their feet. But you overhear someone sitting in the next row complaining, "They've changed the key to C—I can't believe they did that, it sounds so wrong!" The annoyed concertgoer in front of you probably has AP. This entry describes the nature of absolute pitch, its acquisition, and current controversies.

A common misconception is that people with AP are more skilled at being able to notice when a tone is played or sung out of tune, or at being able to sing perfectly intonated tones. Although such differences certainly exist among people, this is a separate phenomenon and not associated with AP. (The fact that there is no agreed-upon name for superior "in tuneness" skills may feed the confusion.) What AP possessors do is notice the pitches (and by extension, the keys) of pieces of music, something that most of us do not. Sing "Happy Birthday" at a birthday party three times, and you may well be singing it in three different keys, and not be at all bothered by it. This is because the identity of a melody is defined by the *intervals,* not by the actual tones used.

Part of a long-standing fascination with AP is that it is so rare (some estimates put it at fewer than 1 in 10,000 people) and that among musicians it is often regarded as the ultimate in musical endowment. Although nearly all musicians can name tones if they are given a reference tone to start with (using their knowledge of intervals to calculate tone names relative to this reference), those with true AP can do so without any external referent, relying instead on an internal template. Some musicians have acquired a stable, long-term memory representation for a single musical tone (typically their tuning note) and can use their knowledge of intervals to "calculate" the names of other tones; this is called quasi-absolute pitch. For this reason, an important methodological consideration in screening for AP possession is that reaction times must be collected.

Why do so many more of us learn to name colors than pitches? Tone labels lack the ecological salience that color labels do, most parents don't teach them, and appreciation of music generally requires that one focuses attention on patterns, intervals, and relations, rather than on the actual pitches of tones. The preponderance of evidence suggests that AP must be acquired early in life, generally before the ages of 8 to 12. There is no documented case of an adult acquiring true absolute pitch. What seems necessary is the early and systematic pairing of tone sensations with verbal labels, whether acquired through explicit or implicit exposure. Gottfried Schlaug found enlargement of the planum temporale in musicians with AP, but the implications of this remain unclear.

Robert Zatorre has shown that similar regions of the dorsolateral prefrontal cortex are activated when musicians with AP name tones, and when musicians with only relative pitch name intervals; this strongly suggests that AP is fundamentally a cognitive operation (of labeling) rather than a perceptual operation (of improved discrimination).

Geneticists have found familial clusters of AP ability, providing evidence that there may exist an allele underlying AP ability. Skeptics counter that AP most certainly has to be learned (no infant automatically knows the names of notes, nor the association between these names and the perceptual input of particular sounds), and therefore even if AP was an entirely learned ability, it would still run in families because only those parents who have it or value it would teach it to their children.

A current controversy concerns the distribution of AP ability across the general population. An online study of 2,000 individuals with AP found the ability to be bimodal: Some people had it, others didn't, and there were no respondents found to possess only partial AP. Other studies have found that the ability is distributed along a continuum. For example, some individuals score 75 or 80% correct on AP tests, indicating that they have AP for only some of the tones in the scale, or that they are performing well above chance, but inconsistently for reasons not yet known. Latent AP has been demonstrated even in nonmusicians who tend to sing their favorite pop songs in the correct key, leading to a two-component model of AP: that it consists of pitch memory separate from pitch labeling.

Another controversy concerns the extent to which infants have a pre-linguistic form of AP, and research findings are contradictory. Diana Deutsch has argued that AP is more prevalent among speakers of tone languages (such as Mandarin), but methodological questions concerning appropriate experimental controls have made this claim controversial. Finally, anecdotal claims that special populations (e.g., individuals with autism or Williams syndrome) show a higher prevalence of AP are not yet widely accepted by the scientific community.

Daniel Levitin

See also Audition: Pitch Perception; Individual Differences in Perception; Perceptual Learning

Further Readings

Athos, E. A., Levinson, B., Kistler, A., Zemansky, J., Bostrom, A., Freimer, N., & Gitschier, J. (2007). *Proceedings of the National Academy of Sciences USA, 104,* 14795–14800.

Deutsch, D. (2006). The enigma of absolute pitch. *Acoustics Today, 2*(44), 11–19.

Levitin, D. J., & Rogers, S. E. (2005). Absolute pitch: Perception, coding, and controversies. *Trends in Cognitive Sciences, 9*(1), 26–33.

Zatorre, R. J. (2003). Absolute pitch: A model for understanding the influence of genes and development on neural and cognitive function. *Nature Neuroscience, 6,* 692–695.

ACOUSTICS AND CONCERT HALLS

When sound travels from a source to a listener in a room, the environment will shape the sound considerably, mainly through the reflections from boundaries. There are a number of qualities of received sound, and each is linked to the characteristics of sound reflections in certain ways. Although in the evolution of Greek and Roman theatres, the acoustics was already considered and gradually improved with increased reverberation and loudness, the subject of architectural acoustics was only established after W. C. Sabine's pioneering study on reverberation early in the 20th century. He carried out quantitative measurements of reverberation, and it was then possible to relate objective qualities of sound to perception.

The perception of music in a concert hall is multidimensional. In the past few decades, a large number of subjective attributes have been introduced for the acoustic quality of concert halls, mainly through questionnaires and interviews. Some of those attributes are interrelated, and some could be combined or subdivided to form more attributes. Correspondingly, to design concert hall acoustics, several objective indices have been developed. This entry examines the meaning of some subjective attributes and their relationships with the objective indices.

Sound Field

Imagine an acoustic impulse is generated in a hall. The first sound to arrive at the listener will be the sound that travels in a straight line from the source, namely the direct sound. This is followed by a series of sounds that have traveled by paths including one or more reflections from room surfaces. The reflection could be specular, namely the incidence angle equaling the reflection angle, if a surface is large and acoustically smooth (i.e., surface roughness significantly smaller than the wavelengths), or diffuse if the surface is irregular. Compared with the direct sound, the amplitude of a reflected sound is always less because part of the sound energy is absorbed by the reflected surfaces and it travels farther. Reflections that arrive at the listener immediately after the direct sound are called *early reflections*. The late reflections are called the *reverberant sound*.

Reverberance

Reverberance is arguably the most important attribute for the acoustic perception of concert halls. It can be measured by the *reverberation time* (RT), which is defined as the time taken for a sound to decay 60 decibels after a continuous sound source is stopped, although actually it is obtained from the 5- to 35-decibel decay and then extrapolated to correspond to a 60-decibel decay because a signal to background noise ratio of 60 decibels is difficult to achieve in practice. It is a function of the volume of the room and of the amount of sound absorption within it. The perception of reverberance is well correlated to the early slope of a sound decay, so the early decay time (EDT) is often used. EDT is obtained from the initial 10 decibels of the decay slope, and then multiplied by a factor of 6. In an ideal diffuse sound field, the decay curves are perfectly linear, and thus, EDT equals RT.

Reverberation is related to liveliness. With a longer reverberation, especially at middle (e.g., 500 Hz–1 kHz) and high frequencies, people feel a concert hall is more lively, and a hall with a short reverberation is perceived as deadly or dry. If a hall is reverberant at low frequencies, say below 350 Hz, it will sound "warm." At middle frequencies, the recommended occupied reverberation time is 1.5 seconds for chamber music, 2 to 2.4 seconds for symphony music, and more than 2.5 seconds for organ music.

Clarity

To enable musical details to be appreciated, a good clarity, or definition, is required. The clarity is usually measured by C, namely the ratio of early sounds (direct sound and early reflections) to the energy in the later reflections. The early reflections are normally defined as those arrived within 80 milliseconds of the direct sound because the ear will perceive the reflections within 50 to 80 milliseconds and the direct sound as one sound. A similar objective measure of clarity is the center time, which is essentially the center of gravity along the time axis of the squared impulse response, where the use of continuous time weighting avoids a sharp time limit between early and late energy. The speed of music (i.e., tempo) also affects the clarity in a concert hall. The subjective feeling of texture, such as something missing in the music, probably relates to the number and nature of early reflections.

Intimacy

Acoustic intimacy is related to the perception of the space size where the music is performed, indicating whether the listener feels acoustically involved or detached from the music. If the music played in a concert hall gives the impression of being played in a small hall, this hall is regarded to have acoustic intimacy, particularly for music originally associated with small halls such as chamber music. Although it is likely that intimacy depends on a number of objective measures, it seems that when the initial time delay gap (ITDG)—namely the time interval between the arrival of the direct sound and the first reflection—is more than 45 milliseconds, the hall has no intimacy.

Spaciousness

The spaciousness of a concert hall has two contributing components: the apparent source width (ASW), which is the impression that the music is from a source wider than the visual width of actual source, and the listener envelopment (LEV), which is the impression of being surrounded by the reverberant sound field. The former increases when there are more early lateral reflections, and the latter will be stronger when the level of late lateral reflection is high. The apparent source width can be measured by the lateral energy fraction (LEF), which is the ratio of lateral sound energy recorded using a "figure-of-eight" microphone (i.e., receiving sounds from the front and back) compared with that recorded using an omnidirectional microphone (i.e., receiving sounds from all directions), all within 80 milliseconds after the direct sound. For the listener envelopment, the late lateral strength (LG) is an objective measure, which relates the lateral energy arriving after 80 milliseconds to the direct sound at 10 meters from the source. Although at the two ears, the lateral reflections could be different, the interaural cross-correlation coefficient (IACC) is used to evaluate spaciousness through measuring their signal dissimilarity, with the listener facing the performing entity.

Loudness

Although the preferred loudness depends on the type of music and many other factors, an experiment with a piece of orchestral music showed that about 60% of the listeners preferred a listening level of 80 to 85 decibels (A-weighted), and others preferred lower or higher levels. An objective measure of loudness in a concert hall is the strength factor (G), which is the difference between the sound level in the hall and that in a free field at a source-receiver distance of 10 meters, using an identical sound source. In an ideal diffuse sound field, where the sound level is uniform across a hall, the strength depends only on reverberation time and hall volume, but in actual halls, the loudness also depends on the source-receiver distance and reflection patterns. In most existing concert halls, the averaged strength in the seating area is about 3 to 5 decibels at middle frequencies, which is a clearly perceivable level.

Timbre

The perception of music in a concert hall in timbre or tone color, such as warm, bright, or harsh, can be related objectively, to a certain degree, to the room influences on the balance between low, middle, and high frequencies. The bass ratio (BR) is the ratio of low (125 and 250 Hz) to middle (500 Hz and 1 kHz) frequency reverberation time, and the treble ratio (TR) is the ratio of high (2 and 4 kHz) to middle (500 Hz and 1 kHz) frequency reverberation time. Although BR and TR are associated with warmth and brightness, respectively, no clear relationship exists between perceptual dimensions and the objective measures. Similarly, the BR and TR could be based on the early delay time or strength values.

Ensemble

The acoustic quality of a concert hall should be judged from the viewpoints of performers as well. Ensemble is a subjective measure of the ability of the musicians to hear their colleagues in an orchestra, and it is strongly affected by the amount of sound reflections. An objective indicator of ensemble is the support factor (ST), namely the difference between the direct sound and reflections, measured on a stage or in a pit, with a source-receiver distance of 1 meter. Two support factors are used, one considering reflections between 20 to 100 milliseconds, called early support, and the other considering reflections between 100 to 1,000 milliseconds, called late support. A concert hall should also give performers an appropriate feeling of immediacy of response (attack), which is mainly determined by the early reflections from the hall to the performer's ears. Moreover, various sections of an orchestra should be balanced and different instrumental groups should be well blended, thus perceived as being tightly coupled as an integrated body.

Acoustic Defects

An echo can be perceived when a reflection arrives at least 50 milliseconds after the direct sound and its amplitude is considerably higher than that of the reflections arriving in the neighboring times. Flutter echoes often occur between two parallel, acoustically smooth, and reflective walls, where a characteristic twang is created by regularly repeated reflections. Sound focus occurs when concave surfaces reflect sound to certain areas, causing high loudness but leaving others with too little reflected energy. Tonal distortion, which makes the perceived sounds different from the originals as a result of the strengthening or weakening of certain frequencies, could be caused by selective sound absorption by room surfaces or seating areas, or resonance of certain architectural elements such as suspended panels. Acoustic glare refers to early reflections of brittle or hard or harsh quality, often caused by large, acoustically smooth reflectors. Finally, it is important to control background noise level, such as from external noise and vibration sources, ventilation systems, and audience movement, to a required level, generally specified using noise criterion (NC) or noise rating (NR) values.

Design

Conventionally, the shoebox shape has been often used for concert halls, but successful examples can also be found with other forms including fan-shaped, vineyard, and arena. Normally, the seat number should be limited to 3,000 and the maximum seat-stage distance 40 meters for concert halls. Other key design considerations include ceiling height, volume per seat, seating layout and arrangement in section, balcony configurations, suspended reflectors, surface conditions in absorption and diffusion, and stage size, layout, and materials. The absorption of seating areas should be carefully designed, so that the sound quality under occupied and unoccupied conditions is similar, which is important for rehearsal.

The use of computer simulation techniques to predict various objective indices in concert halls started from the 1960s. Auralisation techniques recently have been developed to listen to music in three dimensions at the design stage as if it is played in the actual hall, so that the design can be adjusted if needed. Physical models, usually at a scale of 1:8 to 1:50, are also often used.

Ranking

Unlike speech, for which the quality can be assessed by simply conducting articulation tests, the perception of music in a concert hall depends on multiple attributes, and there are often low mutual correlations between those attributes. For a given attribute, people's preference could vary significantly, although the normal distribution is usually followed. Moreover, the importance of different attributes varies considerably between listeners. Furthermore, the variation in objective indices at various seats in a hall sometimes exceeds that between different halls. Although all these factors make it difficult to exactly rank every concert hall, there is a general consensus about which halls have excellent acoustics, based on subjective attributes or objectives indices.

Jian Kang

See also Audition; Audition: Cognitive Influences; Audition: Disorders; Audition: Loudness; Audition: Pitch Perception; Audition: Temporal Factors; Auditory Illusions; Auditory Imagery; Auditory Scene Analysis; Auditory Thresholds; Sound Reproduction and Perception

Further Readings

Ando, Y. (1985). *Concert hall acoustics.* Berlin: Springer.

Barron, M. (1993). *Auditorium acoustics and architectural design.* London: E & F Spon.

Beranek, L. (1996). *Concert and opera halls: How they sound.* Woodbury, NY: Acoustical Society of America.

Chourmouziadou, K., & Kang, J. (2008). Acoustic evolution of ancient Greek and Roman theatres. *Applied Acoustics, 69,* 514–529.

Cremer, L., & Müller, H. (1978). *Die wissenschaftlichen Grundlagen der Raumakustic* (Vol. 1). Stuttgart, Germany: Hirzel.

International Organization for Standardization (ISO). (1997). *Measurement of the reverberation time of rooms with reference to other acoustical parameters.* Geneva: Author.

Jordan, V. L. (1980). *Acoustical design of concert halls and theatres.* London: Applied Science.

Kuttruff, H. (2000). *Room acoustics* (4th ed.). London: Spon Press.

Rossing, T. (2007). *Handbook of acoustics.* New York: Springer.

Sabine, W. C. (1992). *Collected papers on acoustics.* Los Altos, CA: Peninsula. (Original published in 1922 by Harvard University Press)

ACTION AND VISION

Vision is so fundamental to our experience of the world that we sometimes forget that the visual system also plays a critical role in the control of the movements we make in that world. Consider what happens, for example, when we perform the deceptively simple act of reaching out and picking up our morning cup of coffee. After identifying our cup among all the other objects on the table, we begin to reach out toward the cup, choosing a trajectory that avoids the box of cereal and the glass of orange juice. At the same time, our fingers begin to conform to the shape of the cup's handle well before we make contact. As our fingers curl around the handle, the initial forces we generate to lift the cup are finely tuned to its anticipated weight—and to our predictions about the friction coefficients and compliance of the material from which the cup is made. From the beginning of the movement to the final grasp, vision is the main source of information that the brain uses to carry out the required computations.

Many traditional accounts of vision, though acknowledging the role of vision in motor control, have simply regarded such control as part of a larger function of the visual system—that of constructing an internal model of the external world. In most of these accounts, there is an implicit assumption that, in the end, vision delivers a single representation of the external world—a kind of simulacrum of the real thing that serves as the perceptual foundation for all visually driven thought and action. But during the last two decades, it has become increasingly clear that vision does not deliver a single general-purpose representation of the external world. Instead, it appears that two separate visual systems have evolved in the primate brain, each system specialized for a different purpose. This entry describes those systems, the evidence for them, different metrics and frames of reference, and interactions between the two systems.

Two Visual Systems

One system, *vision-for-perception,* allows us to recognize objects and their relationships, enabling us to build up a knowledge base about the world. This is the system we are more familiar with, the one that gives us our conscious visual experience—and allows us to see and appreciate objects in the world beyond our bodies. The other system, *vision-for-action,* provides the visual control we need to move about and interact with those objects. Vision-for-action does not have to be conscious, but it does have to be quick and accurate.

Although there are many different pathways over which visual information is conveyed from the eyes to the brain, the distinction between vision-for-perception and vision-for-action has been mapped onto two anatomically separate "streams" of visual projections that arise from early visual areas in the cerebral cortex of the primate brain (see Figure 1). The ventral visual stream, which projects to higher-order visual areas in the ventral part of the temporal lobe, mediates vision-for-perception. The dorsal visual stream, which projects to visuomotor areas in the posterior parietal lobe, mediates vision-for-action. Importantly, the dorsal stream also gets visual inputs directly from subcortical visual structures that bypass the early visual areas in the cortex.

The ventral stream, which plays the critical role in vision-for-perception, transforms incoming visual information into perceptual representations that embody the enduring characteristics of objects and their relations. These representations allow us to perceive the world beyond our bodies, to share that experience with other members of our species, and to plan a vast range of different actions with respect to objects and events that we have identified. This constellation of abilities is often identified with consciousness, particularly those aspects of consciousness that have to do with decision making and reflecting on our own thoughts and behavior. The perceptual machinery in the ventral stream that has evolved to do this is not linked directly to specific motor outputs, but instead accesses action plans via cognitive systems that rely on memory, semantics,

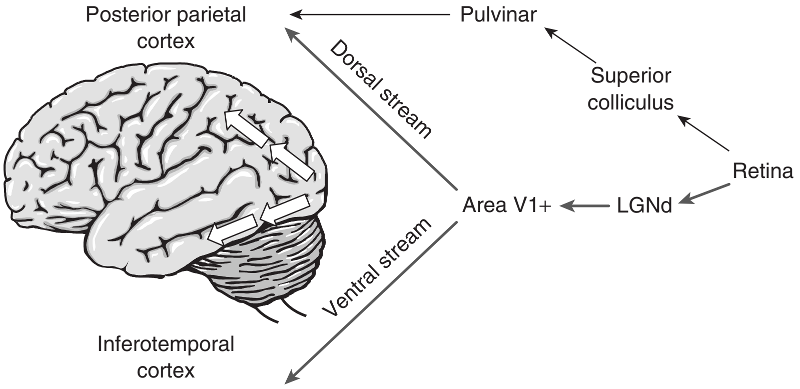


Figure 1 Schematic Representation of the Two Streams of Visual Processing in the Human Cerebral Cortex

Notes: The retina sends projections to the dorsal part of the lateral geniculate nucleus in the thalamus (LGNd), which projects in turn to the primary visual cortex in the occipital lobe of the cerebral cortex. Within the cerebral cortex, the ventral stream arises from early visual areas (V1+) and projects to regions in the ventral parts of the temporal cortex. The dorsal stream also arises from early visual areas but projects instead to the posterior parietal cortex. The posterior parietal cortex also receives visual input from the superior colliculus via the pulvinar, a nucleus in the thalamus. On the left, the approximate locations of the pathways are shown on an image of the brain. The routes indicated by the arrows involve a series of complex interconnections.

spatial reasoning, and communication with others. Once a course of action has been selected, however, the visual control of the constituent movements of that action is mediated by the dorsal stream, whose visuomotor modules are intimately linked with motor areas in the cerebral cortex and other (older) motor systems in the basal ganglia and brainstem. Thus, the structural and spatial attributes of a goal object are processed by both streams, but for different purposes. This is not to say that the distribution of visual inputs does not differ between the two streams but, rather, that the main difference lies in the nature of the transformations that each stream performs on those two sets of inputs. In the case of the ventral stream, information about a broad range of object parameters is transformed for perceptual purposes; in the case of the dorsal stream, many of these same object parameters are transformed for the control of actions. But even though the two systems transform visual information in quite different ways, they work together in the production of adaptive behavior. In general terms, one could say that the selection of appropriate goal objects depends on the perceptual machinery of the ventral stream, whereas the visual control of the goal-directed action is carried out by dedicated online control systems in the dorsal stream.

Neuropsychological Evidence

Much of the evidence for the distinction between vision-for-action and vision-for-perception comes from work with neurological patients. One of the most compelling cases is that of D. F., a young woman whose ventral stream was selectively damaged as a consequence of hypoxia from carbon monoxide poisoning. D. F. shows no visual awareness of the form and dimensions of objects (visual form agnosia) and cannot recognize common objects or the faces of her friends and relatives. Indeed, D. F.'s deficit is so profound that she cannot discriminate between simple shapes such as a triangle and square, and cannot tell a horizontal from a vertical line. Brain imaging has revealed that she has bilateral damage in that part of the ventral stream that is involved in processing the geometrical structure of objects. Nevertheless, when D. F. reaches out to grasp objects, the posture of her hand and fingers is exquisitely tuned in flight to the size, shape, and orientation of the object in front of her, just as it is in a person with normal vision. For example, when she is presented with a series of rectangular blocks that vary in their dimensions but not in their overall surface area, she is unable to say whether or not any two of these blocks are the same or different. Even when a single block is placed in front of her, she is unable to indicate how wide the block is by opening her index finger and thumb a matching amount. Nevertheless, when she reaches out to pick up the block using a precision grip, the opening between her index finger and thumb is scaled in flight to the width of the block, just as it is in people with normal vision. Similarly, even though she cannot indicate the orientation of elongated objects placed in a table in front of her, she rotates her hand appropriately when she reaches out to grasp the same objects.

Patients who have suffered selective damage to the dorsal stream show a pattern of deficits and spared abilities that are the mirror image of those seen in D. F. Such individuals are unable to use visual information to guide their hand movements as they reach out to grasp objects, even though they can see the object and can describe where it is and what it looks like. At the same time, these patients typically have no difficulty using input from other sensory systems, such as touch or audition, to guide their movements. In other words, their deficit is neither "purely" visual nor "purely" motor but is instead a visuomotor deficit. The clinical name for this neurological disorder is *optic ataxia*.

Different Metrics and Frames of Reference

The evidence from neurological patients points to a clear dissociation between the visual pathways supporting perception and action. But why did two separate streams of visual processing evolve in the primate brain? Or, to put it another way, why couldn't one general-purpose visual system handle both vision-for-perception and vision-for-action? The answer to this question lies in the different computational requirements of the two kinds of vision. Simply put, perception and action require quite different transformations of the visual signals. To be able to grasp an object successfully, for example, the brain must compute the actual size of the object, and its orientation and position with respect to the grasping hand of the observer (i.e., in egocentric coordinates).

The time at which these computations are performed is equally critical. Observers and goal objects rarely stay in a static relationship with one another and, as a consequence, the egocentric coordinates of a target object can often change radically from moment to moment. For these reasons, it is essential that the required coordinates for action be computed in an egocentric framework at the moment the movements are to be performed. In other words, vision-for-action works largely in an "online" mode. Unlike our visual percepts, the visual information that is used to control actions such as grasping is not accessible to conscious scrutiny.

Perceptual processing needs to proceed in a quite different way. Vision-for-perception does not require the absolute size of objects or their egocentric locations to be computed. In fact, such computations would be counterproductive because we almost never stay fixed in one place in the world. For this reason, it would be better to encode the size, orientation, and location of objects relative to each other. Such a scene-based frame of reference permits a perceptual representation of objects that transcends particular viewpoints, while preserving information about spatial relationships (as well as relative size and orientation) as the observer moves around. Indeed, if the perceptual machinery had to deliver the real size and distance of all the objects in the visual array, the computational load would be prohibitive. It is far more economical to compute just the relational metrics of the scene, and even these computations do not always need to be precise. The reliance on scene-based frames of reference means, for example, that we can watch the same scene unfold on a small television or on a gigantic movie screen without being confused by the changes in scale. These differences in the frames of reference and metrics used by vision-for-perception and vision-for-action can give rise to striking differences in the sensitivity of the two systems to visual illusions. Thus, the scaling of grasping movements is often insensitive to size-contrast illusions, which (by definition) affect our perception of object size (see Figure 2).

Evidence also suggests that our perception of an event takes longer to emerge than does our ability to act upon that event. Thus, we are often able to react to the change in the trajectory of a moving target well before we perceive that change.

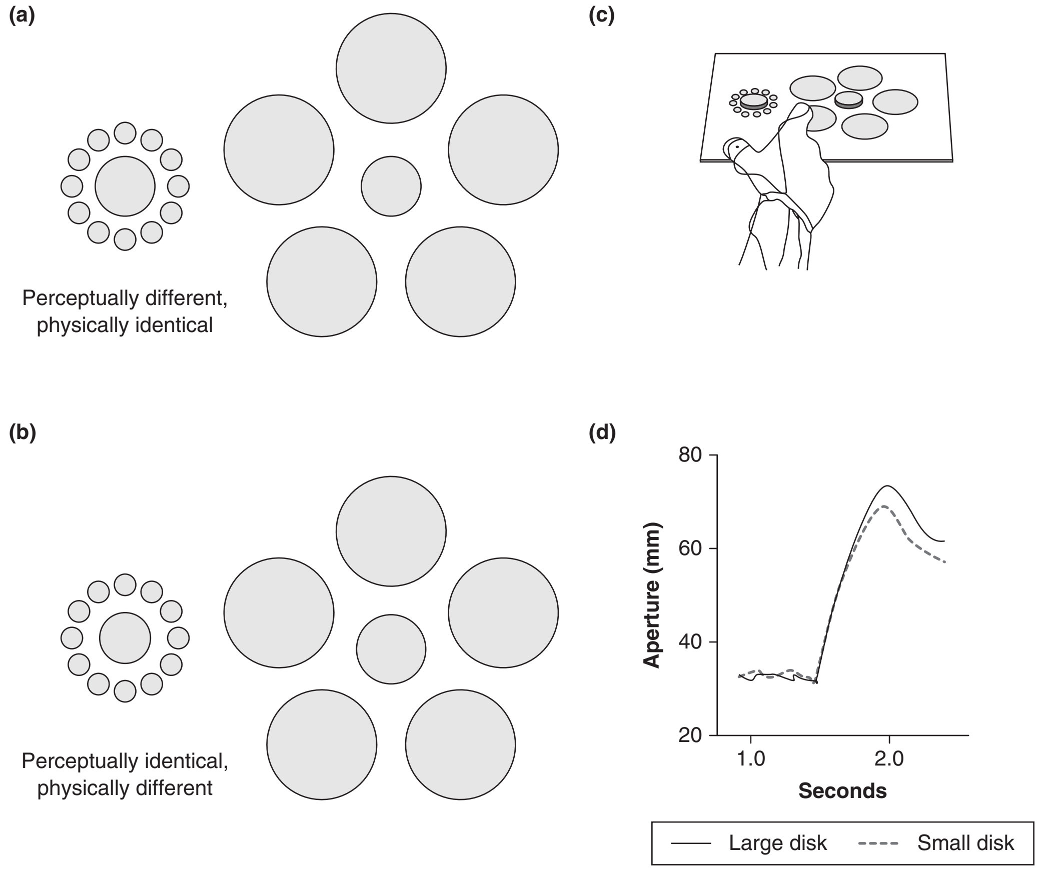


Figure 2 The Effect of a Size-Contrast Illusion on Perception and Action

Source: Adapted from Aglioti, S., DeSouza, J., & Goodale, M. A. (1995). Size-contrast illusions deceive the eyes but not the hand. *Current Biology, 5,* 679–685.

Notes: (a) The traditional Ebbinghaus illusion in which the central circle in the annulus of larger circles is typically seen as smaller than the central circle in the annulus of smaller circles, even though both central circles are actually the same size. (b) The same display, except that the central circle in the annulus of larger circles has been made slightly larger. As a consequence, the two central circles now appear to be the same size. (c) A 3-D version of the Ebbinghaus illusion. People are instructed to pick up one of the two 3-D disks placed either on the display shown in panel A or the display shown in panel B. (d) Two trials with the display shown in panel B, in which the person picked up the small disk on one trial and the large disk on another. Even though the two central disks were perceived as being the same size, the grip aperture in flight reflected the real rather than the apparent size of the disks.

Moreover, the products of perception also need to be available over a much longer time scale than does the visual information used in the control of action. We may need to recognize objects we have seen minutes, hours, days—or even years before. To achieve this, the coding of the visual information has to be somewhat abstract, transcending particular viewpoint and viewing conditions. By working with perceptual representations that are object- or scene-based, we are able to maintain the constancies of size, shape, color, lightness, and relative location, over time and across different viewing conditions. Although there is much debate about the way in which this information is coded, the identity of the object and its location within the scene, rather than its disposition with respect to the observer is of primary concern to the perceptual system. Thus, current perception combined with stored information about previously encountered objects facilitates the recognition of objects and contributes to the control of goal-directed movements when we are working in off-line mode (i.e., on the basis of our memory of goal objects and their location in the world).

Interactions Between the Two Streams of Visual Processing

So let us see how the two streams of visual processing work together in the example introduced earlier: the (deceptively) simple act of picking up a cup of coffee. The initial identification of our cup among the other objects on the table (including the cups of other people) depends on processing in the ventral stream. Areas in this ventral "perception" stream are undoubtedly recruited for identifying the cup—and the handle on the cup. But by virtue of identifying the handle as the goal for the intended action, the ventral stream (together with associated cognitive machinery in other brain regions) also determines the functional grasp to be used. But once the cup has been identified and the intended grasp posture selected, the job of using visual information to program and control the constituent movements for that grasp are carried out by dedicated visuomotor modules in the dorsal "action" stream. These dorsal-stream modules compute the real size of the handle and its location and disposition with respect to our hands and fingers so that our grasp is directed toward the handle in the correct orientation and the opening of fingers is scaled to its size. Moreover, in most cases, all of this is done "automatically" without any conscious supervision.

But what happens when the hand actually reaches its goal? Once the hand and fingers make contact with the handle, the initial grip and load forces required to pick up the cup are exquisitely well matched to the cup's weight, compliance, and surface friction—and to the amount of coffee left in the cup. (When we get that last computation wrong, as we sometimes do, we can end up spilling our coffee all over the table!) The calculation of these forces depends on learned associations between the material properties of the cup and its appearance (as well as our updated memory of its contents). It seems likely that areas in the ventral stream that process object shape and surface cues play a critical role in recognizing the cues associated with the material properties of the cup and other objects we interact with. This leaves unanswered how the ventral and dorsal streams coordinate their computations. Whatever complex interactions might be involved, goal-directed action is unlikely to be mediated by a simple serial processing system; instead, multiple iterative processing across many levels of the brain is almost certainly required.

Final Thoughts

The idea of two visual systems in a single brain might seem initially counterintuitive. Our visual experience of the world is so compelling that it is hard to believe that some other quite independent set of visual information is guiding our movements. After all, it seems obvious that the perceptual representation that allows us to recognize our coffee cup is the same one that enables us to pick it up. But evidence from observations of neurological patients and other studies suggests that this belief is an illusion. The visual processes that lead to our perception of the cup are *not* the same as those that guide our hand as we reach out and pick up it up! The illusion arises because the former but not the latter are associated with conscious visual experience.

Melvyn Alan Goodale

See also Agnosia: Visual; Optic Ataxia; Reaching and Grasping; Visual Illusions; Visually Guided Actions

Further Readings

Aglioti, S., DeSouza, J., & Goodale, M. A. (1995). Size-contrast illusions deceive the eyes but not the hand. *Current Biology, 5,* 679–685.

Andersen, R. A., & Buneo, C. A. (2003). Sensorimotor integration in posterior parietal cortex. *Advances in Neurology, 93,* 159–177.

Bruno, N., Bernardis, P., & Gentilucci, M. (2008). Visually guided pointing, the Müller-Lyer illusion, and the functional interpretation of the dorsal-ventral split: Conclusions from 33 independent studies. *Neuroscience and Biobehavioral Reviews, 32,* 423–437.

Carey, D. P. (2001). Do action systems resist visual illusions? *Trends in Cognitive Sciences, 5,* 109–113.

De'Sperati, C., & Baud-Bovy, G. (2008). Blind saccades: An asynchrony between seeing and looking. *Journal of Neuroscience, 28,* 4317–4321.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15,* 20–25.

Goodale, M. A., & Milner, A. D. (2004). *Sight unseen: An exploration of conscious and unconscious vision.* Oxford, UK: Oxford University Press.

Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology, 13,* 159–166.

Milner, A. D., & Goodale, M. A. (2006). *The visual brain in action* (2nd ed.). Oxford, UK: Oxford University Press.

AESTHETIC APPRECIATION OF PICTURES

Viewing visual art can be a dramatic, richly rewarding yet wholly personal event. In discussing the aesthetic appreciation of pictures within the bounds of perceptual research, an attempt is made to cross a significant chasm between a detailed cellular and cortical understanding of lower-level visual objective properties such as color and shape, and a less clear understanding of higher-level visual phenomenology or subjective experience. Over the centuries, both the definition and content of an aesthetic experience has meant many things to many people; however, perceptual research has tended to define aesthetic appreciation in terms of some measure of *preference* based on the perceived *beauty* of the image in question. In this respect, perceptual research provides echoes of both David Hume's and Immanuel Kant's approaches to aesthetics in terms of their discussion of taste and beauty, respectively. The factors that are thought to influence an individual's aesthetic response to pictures include both the physical properties of the picture itself that exist "inside the frame," as well as the role of contextual influences such as the labeling and presentation of art that exist "outside the frame." Although perceptual research continues to use group data (or a nomothetic approach) as a way of tackling aesthetics, others have argued that individual data (or an idiographic approach) should serve as the unit of currency if we are to have the fullest appreciation of appreciation. This entry describes how aesthetics are measured, objectivist and subjectivist approaches to aesthetics, and how researchers use these.

Measuring Aesthetics

The origins of empirical aesthetics are generally attributed to Gustav Fechner in his book *Elementary Aesthetics,* whereas Daniel Berlyne is typically credited with the resurgence of interest in linking the scientific method with aesthetics in the 1970s. These early experiments were characterized by assessing individuals' preferences for large sets of artificially constructed stimuli such as polygons that differed according to a number of quantifiable properties labeled *collative* (e.g., complexity), *psychophysical* (e.g., color), and *ecological* (e.g., meaning) variables. According to Berlyne's psychobiological account, aesthetic experience should be higher for intermediate levels of arousal, with arousal being calculated as summed activation of the these features: Many-sided polygons should contain less color than polygons with fewer sides, for example. This early work set the trend for the measurement of aesthetic appreciation using a simple numeric (also referred to as a Likert) scale, where individuals were invited to sort or rate pictures from the least preferred/beautiful (e.g., 1) to the most preferred/beautiful (e.g., 5). Although it is easy to criticize such measures on the basis that individuals may not use the whole range of scores available to them, such subjective measures remain at the heart of perceptual research on aesthetics. Over time, however, self-report measures of aesthetic experience have been augmented with additional objective measures such as looking time and blood oxygenation levels in the brain to provide converging evidence in understanding the aesthetic experience.

Aesthetics Inside the Frame

These first steps in understanding aesthetics via perceptual research reveal what is essentially a reductionist approach. Although not of primary interest to Berlyne, the assumption here is that it should be possible to arrive at an understanding of the genesis of beauty within an actual work of art by studying individuals' responses to the building blocks of visual perception. In this way, preference for certain pictorial images is decomposed into the study of preference for particular conjunctions of colors, line orientations, sizes, and shapes. Common to many psychological domains, there is a tension here between the control that one is able to exert on so-called pictorial art within the laboratory and the ability to generalize this data to the much richer examples of pictorial art found in the real world. On the one hand, the development of self-produced abstract visual stimuli means that individuals will have had no prior exposure with the images. This makes the aesthetic experience a primitive one, where influences of schema or memory are reduced and the image is assessed in a stimulus-driven or bottom-up manner. On the other hand, these types of stimuli are far removed from the actual stimuli of interest and ecological validity is weak: Can the study of polygons really tell us anything about Picasso?

However, an intersection between low-level and high-level visual experience is provided by the work of Piet Mondrian, whose painted elements specifically map onto basic visual features such as line orientation and color. Consequently, researchers have been able to manipulate the line spacing, line thickness, orientation, and placing of color within the painting to assess the extent to which individuals judge an original Mondrian composition to be more aesthetically pleasing than a manipulated Mondrian. Evidence has accrued to suggest that even individuals untrained in the visual arts can give higher ratings to the original paintings, suggesting that variation in aesthetic experience is driven partly by the arrangement of the visual elements presented to the viewer. Other researchers have shown that aesthetic preference for original rather than manipulated paintings also apply to representational artwork, although success in establishing preference for the original image often relies on major rather than minor spatial changes in the piece. These observations suggest that an aesthetically pleasing painting will be one that has established *the* best arrangement (or balance) of elements, and that these preferred arrangements deliberated over by the artist are then readily perceivable by non-artists. These findings align well with the principle of Prägnanz found in Gestalt psychology (also referred to as "visual rightness") and provide evidence for universalism in aesthetic experience.

Aesthetics Outside the Frame

In contrast to an objectivist approach to the empirical study of aesthetics in which the beauty of the picture is thought to be somehow engrained within the organization of the visual elements themselves, subjectivist approaches emphasize the role of external variables in deciding what is beautiful and what is not. The need for a subjectivist component to aesthetics will be clear to anyone who has had the misfortune to accompany a lover of Renaissance painting around a modern art gallery. That individuals can have different responses to identical visual stimulation therefore implies that attitudes toward and experience with art can have significant influences on aesthetic appreciation. Comparisons between art novices and art experts are often employed within the literature, although the determination of what it means to be an "art expert" or "what constitutes an art expert" can range from a median split within a given sample to the achievement of (at least) graduate-level training. On the basis of previous distinctions between figurative and abstract art, and original color or manipulated black-and-white images, art novices' aesthetic preferences tend toward color images of figurative art, whereas art experts tend to have much broader preferences.

The labeling of a picture is also thought to have an influence on an individual's aesthetic response. However, the outcome of labeling depends on both the content of the label and the type of picture to which it is applied. For example, the provision of a descriptive title to a representational picture may be redundant (e.g., William Turner's *The Shipwreck* [1805]), but when applied to a more abstract work (e.g., Marcel Duchamp's *Nude Descending a Staircase* [1912]) such information may help the individual unlock the ambiguous elements on view. Moreover, elaborative comments related to the genesis, style, or interpretation of the piece can also influence an individual's reaction to it. For example, informing the viewer that in Edvard Munch's *The Scream* (1893) the character at the forefront of the painting is not actually screaming but rather trying to protect the self from the scream of Nature, may radically affect the viewer's perception and aesthetic appreciation of the image. Although the effect of labeling does not seem to be consistent for all images, studies which have manipulated the presentation of no label versus descriptive or elaborative labels across artworks have shown that descriptive labels are often useful for increasing the understanding associated with a picture, but elaborative labels consistent with the artwork are more likely to lead to a greater aesthetic reaction.

Other external influences on aesthetic appreciation stem from exactly where the picture is viewed. In the interests of control, individuals taking part in studies on empirical aesthetics are likely to view images on a computer screen for a prescribed period. This is quite different from the natural viewing conditions of an art gallery, in which the original size of the picture is preserved, the viewing distance is often carefully controlled, and the viewing time is essentially self-paced. Consequently, images viewed in aesthetic appreciation studies are likely to be substantially reduced in both size and viewing time, relative to the original viewing of the artwork. Although the limited data that examines the comparison between original and reduced size images reveals little substantive difference between the two forms of presentation, some optical effects or effects of scale intended by the artist could potentially be lost by the manipulation of size to suit ease of presentation in the laboratory. For example, the large color-field paintings of Mark Rothko could be judged quite differently if there were no preservation of the size of the image itself. Limiting the presentation time of a picture from a self-paced average of around half a minute can also restrict the types of pictorial analyses conducted, with short presentation times leading to aesthetic judgments based on only the global properties of the image itself.

Is There Any Account for Taste?

In examining objectivist and subjectivist approaches to the aesthetic appreciation of pictures, a consolidation of both approaches (termed *interactivist*) has begun. In the defense of objectivist models, both representational and nonrepresentational art can elicit aesthetic responses and, as such, the relationship between these judgments should be in the form of the picture rather than its content. In defense of subjectivist models, identical visual stimulation can lead to different aesthetic preferences between individuals. However, it is also necessary to consider alternatives to the nomothetic approach to empirical aesthetics. Even with the paring down of complex pictorial stimuli to their basic-level components, researchers have found it difficult to develop a group model of aesthetic appreciation that adequately represents the individual. Moreover, clinical applications of aesthetics have tended to adopt an idiographic approach. For example, the palliative benefit derived from the display of artwork in healthcare settings may well rest with personal rather than institutional selection. Further, although patients in the early stages of Alzheimer's differ between themselves in their categorization of pictures, aesthetic preferences for the pictures can remain stable over a two-week period whereas their explicit memory does not. Finally, in systematically varying pictorial representations of male and female bodies as a way of evaluating both own and ideal body image, individual aesthetic judgments of the images depend heavily on a number of psychosocial factors at the time of selection. Understanding aesthetics at either an individual or a group level promises the tailoring of more intense and more beautiful experiences in our environment. The study of empirical aesthetics has shown that it is indeed possible to begin to account for taste, although some of the most crucial aspects of the aesthetic experience remain at a distance.

Ben Dyson

See also Attractiveness; Color Perception; Context Effects in Perception; Gestalt Approach; Pictorial Depiction and Perception; Top-Down and Bottom-Up Processing

Further Readings

Berlyne, D. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century Crofts.

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology, 95,* 489–508.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review, 8,* 364–382.

Redies, C. (2007). A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spatial Vision, 21,* 97–117.

AFTEREFFECTS

A perceptual aftereffect can be said to occur when an inducing stimulus (typically presented for a prolonged period) affects subsequent perception of another stimulus. As described in this entry, aftereffects are a seemingly maladaptive consequence of the processes of sensory adaptation by which our sensory systems are continually matching their response properties to the prevailing environment. In the normal course of events, our perception of the world does a remarkable job of "telling us what's there," and it is hard to appreciate just how much work our sensory systems are doing to support this feat. Aftereffects represent instances when there is a clear mismatch between our perception and our surroundings and can thus provide insight into the mechanisms underlying our perceptual experience.

As a specific example, imagine that you have just turned your car onto a busy highway. As you accelerate up to speed, the world outside seems to be rushing past at an alarming rate. As you drive along, your perception quickly adapts to this high-speed environment and speeds that just minutes ago seemed alarmingly fast are now the new norm. However, you must be careful when you leave the highway and return to the ordinary road because your adaptation to fast highway speeds can have a pernicious aftereffect. Specifically, adaptation to high speeds causes an aftereffect in perceived speed whereby subsequent slow speeds appear even slower than they should, potentially leading you to drive too fast after leaving the highway and to underestimate the time required for braking.

In addition to affecting the perception of speed, adaptation to visual motion can also cause static objects to appear in motion. Consider what happens, for example, after prolonged viewing of water rushing down a waterfall. If you transfer your gaze from the falling water to the nearby rocks, you can experience the illusion that the rocks are moving in the opposite, upward, direction: the *waterfall illusion.* Paradoxically, this illusory perception of motion entails little or no change in the apparent position of the rocks; although the rocks appear to be moving upward, they never seem to get to a new position. This dissociation between the perception of motion and location during the waterfall illusion indicates that it is a consequence of adaptation of sensory mechanisms specific to the processing of motion.

Analogous Aftereffects in Vision, Touch, and Hearing

The waterfall illusion has analogues both in hearing and in touch. In all these sensory modalities, prolonged exposure to a moving stimulus biases subsequent perception in the opposite direction: a "negative" aftereffect. The auditory analogue, for example, can be induced by prolonged exposure to a sound source sweeping horizontally from left to right. A subsequently presented static sound is then perceived to move in the opposite direction, right to left. The existence of analogous aftereffects across different sensory modalities is evidence that they reflect common strategies of sensory processing. A motion aftereffect has even been reported across modalities, such that after a few minutes of viewing a square moving in depth, a steady sound is perceived as changing loudness in the opposite direction. For example, after prolonged viewing of a square increasing in size to simulate motion toward the observer, a subsequent steady test tone can be heard to decrease in loudness as though it were moving away from the observer.

Adaptation and its consequent aftereffects are by no means restricted to motion. Actually, they are nearly ubiquitous in sensory processing. For example, after a period of adaptation to a tilted grating, an objectively vertical grating appears to take on a marked tilt opposite to the tilt of the adapting grating. You can experience the tilt aftereffect yourself by following the instructions in the legend to Figure 1.

How Can We Tell Where Aftereffects Are Generated?

Staring at a bright light can saturate the photoreceptors in our retinas, leading to the perception of afterimages. Perceptual afterimages are characteristically monocular. This means that after staring at a bright light with the right eye covered or closed, the afterimage can only be seen with the left eye and does not transfer to the right. However, aftereffects produced by adaptation of neurons in the visual cortex, where signals from the two eyes converge, typically transfer from one

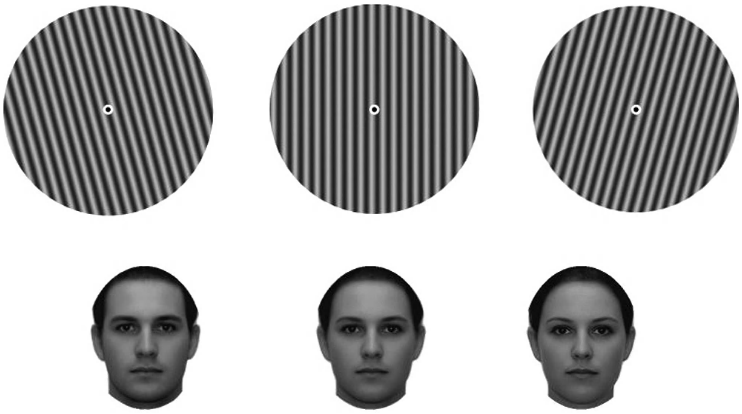


Figure 1 Experiencing Aftereffects

Notes: Experience aftereffects for yourself using these images: (Top) The tilt aftereffect; (bottom) an aftereffect of facial gender. *Tilt aftereffect:* The procedure for inducing a tilt aftereffect is as follows. First, satisfy yourself that the middle grating stimulus is oriented vertically. Then transfer your gaze to the fixation marker at the center of one of the tilted gratings. Count to 50 while keeping your eyes on the fixation marker. There is no need to stare too fixedly or the light and dark bars of the grating will induce differential adaptation at the retina leading to an afterimage. Instead, just let your eyes wander a little over the white annulus of the fixation marker. You might want to cover the other gratings during this adaptation phase as they can be a little distracting but it is not necessary for getting the effect. Once you have counted to 50, transfer your gaze to the center of the middle grating. It should now appear to be tilted away from vertical in the direction opposite to the adapting grating—a tilt aftereffect! If you now adapt to the other tilted grating you can reverse the effect. Adapting to one of the tilted gratings with one eye covered then looking at the middle grating with the other eye should still produce an aftereffect. This inter-ocular transfer of the tilt aftereffect indicates that it is mediated by adaptation of neurons in the visual cortex, where signals from the two eyes converge, rather than earlier in the visual pathway. *Gender aftereffect:* The procedure for inducing an aftereffect of facial gender is as follows. First, look at the middle, androgynous face and decide whether it looks male or female to you. If it looks male, you should adapt first to the male face to get the best effect. If it looks female to you, adapt to the female face. To adapt, look at the appropriate face with the other two faces covered and count to 50. Once you have counted to 50, uncover the middle face. It should appear to have changed gender! If you now adapt to the other face, the gender of the middle face should appear to change back.

eye to the other. For example, the magnitude of the tilt aftereffect is only slightly reduced by presenting adaptor and test in different eyes—you can see this for yourself by putting your hand over your right eye while adapting to one of the oriented gratings in Figure 1, then moving your hand to your left eye to view the objectively vertical middle grating.

Behavioral studies with human observers have demonstrated high-level perceptual aftereffects to properties of faces such as their identity, gender, ethnicity, emotion, and attractiveness. For example, after prolonged viewing of a male face or set of male faces, a face that previously appeared androgynous (gender neutral) will now be perceived as opposite in gender to the inducer(s), that is, female. These effects transfer to a large extent between faces of different size, position, and orientation. This invariance to low-level image properties indicates that face aftereffects are generated at a high level of the visual processing hierarchy. To witness an aftereffect of facial gender, follow the instructions in the legend to Figure 1.

Contingent Aftereffects

The motion aftereffect following prolonged exposure to moving colored (e.g., red) dots appears stronger when the subsequently viewed static dots are the same color (red) than when they are a different color (green). This color selectivity allows the direction of the motion aftereffect to be made contingent on the color of the dots. For example, after prolonged exposure to red dots moving up and green dots moving down, colored static dots will tend to appear to move in the opposite directions: down if they are red and up if they are green.

Similarly, in auditory perception, the motion aftereffect can be made contingent on whether the sound is rising or falling in pitch. For example, an auditory motion aftereffect contingent on direction of pitch change can be induced by prolonged exposure to a rightward-moving sound with falling pitch alternating with a leftward-moving sound rising in pitch. A subsequent stationary test sound with falling pitch can then seem to move to the left, and a stationary test with rising pitch is heard to move rightward.

The existence of contingent aftereffects is an indication of the dynamic nature of the way that neural representations of different sensory attributes can become associated with one another to reflect complex relationships in the environment. A fuller understanding of mechanisms underlying contingent aftereffects should allow us to appreciate how different sensory attributes are bound together to bring about our unified perceptual experience of the world.

Colin W. G. Clifford

See also Afterimages; McCollough Effect; Olfactory Adaptation; Prism Adaptation; Selective Adaptation; Taste Adaptation; Visual Light- and Dark-Adaptation

Further Readings

Clifford, C. W. G., & Rhodes, G. (Eds.). (2005). *Fitting the mind to the world: Aftereffects in high-level vision.* Oxford, UK: Oxford University Press.

Dong, C.-J., Swindale, N. V., & Cynader, M. S. (1999). A contingent aftereffect in the auditory system. *Nature Neuroscience, 2,* 863–865.

Kitagawa, N., & Ichihara, S. (2002). Hearing visual motion in depth. *Nature, 416,* 172–174.

Mather, G., Verstraten, F., & Anstis, S. (Eds.). (1998). *The motion aftereffect: A modern perspective.* Cambridge: MIT Press.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428,* 558–561.

AFTERIMAGES

Nearly everyone has experienced seeing the flash from a camera for several seconds or even minutes after the flash occurred. This is an example of the phenomenon known as an afterimage. Not everyone is aware, however, that it is just one of the more noticeable effects of visual mechanisms that underlie all visual experience.

An afterimage is a spatial pattern that is seen when a stimulus that produced it is no longer present and resembles the stimulus. Its properties depend on the intensity, color, and duration of the stimulus and on the stimuli that follow it. It can last a fraction of a second or several minutes. It takes two forms: *positive,* if the colors and luminances (the light and dark regions) correspond to those in the stimulus, and *negative* if they are opposite. (Opposite colors here are complementary colors, such as red and green or yellow and blue.) These two forms are related to two separate causes of afterimages, *visual persistence* and *adaptation,* which are described in this entry.

Visual Persistence

Visual persistence results from the inability of the nerve cells in the visual system, especially the photoreceptors, to follow changes of stimulation instantaneously: A finite amount of time is required for them to respond to the onset of the stimulus and, more important for present purposes, to respond to its cessation as well. As long as the nerve cells are responding to and therefore signaling the presence of a stimulus, one continues to see it, just as one continues to see the flash of the camera if one's surroundings are otherwise dark or if one closes one's eyes. One can also observe such persistence by placing a white spot on a dark disk and spinning the disk. The spot appears to have a tail, like that of a comet. The faster the disk spins, the longer the tail. If the tail just barely extends all the way around the disk, the time required to complete one full rotation is a measure of the duration of the visual persistence. Visual persistence produces positive afterimages, and this allows us to see the images in movies and television as though they were stable, moving objects instead of a succession of still images.

Adaptation

An animal's survival depends on sensing changes in its environment, and sensitivity to such changes is fostered by keeping sensory excitation at intermediate levels and by damping the response to stimuli that do not change. Negative afterimages are a by-product of the processes that accomplish this.

The process of light adaptation reduces the effect of light much as light-sensitive sunglasses do. Suppose that the light level increases tenfold, and that as a consequence, the sunglasses darken so that only 30% of the light passes through. Then a tenfold increase in light ends up increasing excitation only threefold. Light adaptation accomplishes the same thing as darkening the sunglasses, except that the change in effectiveness of the light occurs not by absorption but by changes in the physiology of the cells of the retina, the thin layer of light sensitive receptor cells, and associated nerve cells on the inside surface of the eye.

However, light affects the sensitivity of only the receptor cells that absorb it and the neurons downstream in the visual pathway to which they are connected. If the camera flash mentioned previously temporarily reduces the effectiveness of light falling on the receptors stimulated by the flash, then during subsequent viewing of a uniform surface such as a white wall, the effectiveness of the light falling on the array of receptors will not be uniform, and the effect of the light falling on the receptors previously stimulated will be reduced. As such, light is less effective; it has the same effect as if a dark pattern on the wall were imaged specifically on the adapted receptors, and so that is what one sees: a dark pattern in the shape of the camera flash. The pattern is dark instead of light, so it is called a *negative afterimage.*

Color

If more of the light in the stimulus is absorbed by the photoreceptors sensitive to red light than, say, the photoreceptors sensitive to green light, then the red-sensitive receptors will be more adapted (desensitized) than the green-sensitive receptors will. Subsequent exposure to white light, which contains a mixture of all colors, will excite the green-sensitive receptors more than the red-sensitive receptors, and so the white light has the same effect as it would have if the light had more green light than red light; consequently, in the areas of the retina adapted to red light, white light looks green. As the resulting color appearance is approximately that of the complement of the color to which the retina is adapted, it also is called a negative afterimage.

Temporal Factors

Even a weak stimulus, if viewed steadily for several seconds or a minute, can accumulate strong, negative afterimages. Conversely, even an extremely brief stimulus, if intense enough, can create strong positive or negative afterimages. These images can last for several minutes, although they pass through alternating phases of disappearance and reappearance. This is partly because maintaining visibility of an image requires movement between the image and the retina. Even when one tries to fixate steadily on one spot, the eye makes small, involuntary movements that ordinarily are both necessary and sufficient to keep images visible, but afterimages move with the retina and thwart the sustaining effect that the eye movements would otherwise have. Blinking or other changes in stimulation usually restores visibility of the afterimage. Long-lasting afterimages also pass through a succession of intense colors, called the *flight of colors,* even when the stimulus itself was white. Weak afterimages that would normally last only a few seconds can be permanently obliterated by brief superimposition of a pattern rich in contours.

Walter Makous

See also Color Perception; Color Perception: Physiological; Retinal Anatomy; Visual Light- and Dark-Adaptation

Further Readings

Brown, J. L. (1965). Afterimages. In C. H. Graham (Ed.), *Vision and visual perception* (pp. 479–503). New York: Wiley.

Lu, H. J., Zavagno, D., & Liu, Z. L. (2006). The glare effect does not give rise to a longer-lasting afterimage. *Perception, 35*(5), 701–707.

Wade, N. J. (1998). *A natural history of vision.* Cambridge: MIT Press.

AGEING AND CHEMICAL SENSES

Reading glasses and hearing aids are accepted accoutrements of the elderly: It is common knowledge that the visual and auditory senses dim with age. What about the chemical senses of smell and taste? A decline in the ability to appreciate ("taste") food and drink in old age was recognized thousands of years ago and noted in the Old Testament of the Bible (2 Samuel 19:34–35). But in what ways do our perceptions of the volatile and water-soluble molecules we smell and taste change with age? How pronounced is that change? And how early in life does it begin?

This entry addresses these questions, but it is important to recognize first that smell (olfaction) and taste (gustation) are distinct physiological systems. They have different receptor types and peripheral neural pathways, and they respond, largely, to different types of chemical stimuli. However, central nervous system responses to many smell and taste stimuli overlap. Thus, when food is eaten—simultaneously releasing some molecules that stimulate taste receptors in the oral cavity and some that flow from the mouth to the nasal cavity and stimulate olfactory receptors—it is perceptually difficult to distinguish which components of the resulting flavor perception are smells and which are tastes. Actually, however, these two aspects of food flavors are not equally affected by ageing: Diminutions in olfactory sensitivity are more pronounced than are diminutions in gustatory sensitivity.

Olfaction

Olfactory receptors are found on primary olfactory receptor neurons (ORNs) located in a relatively small patch of tissue high in the nasal cavity. These neurons extend cilia along the epithelial surface and, thus, are uniquely exposed to the external environment and subject to a constant barrage of potentially toxic chemicals and particulates, as well as being susceptible to direct injury from microbes. Although ORNs are also highly unusual in that there is ongoing replacement of these neurons throughout life, this is a complex process requiring reinnervation of the olfactory bulb (the first brain relay in the olfactory pathway), and it is often imperfect. Degeneration of the olfactory neural epithelium, and patchy replacement by respiratory epithelium, is seen even in young adults and becomes more pronounced with aging.

Olfactory function is most often assessed via tests of threshold sensitivity (the lowest concentration of an odorant that can be detected) or of the ability to identify suprathreshold concentrations of odors (e.g., is this orange, licorice, grass, or banana?); less frequently, ratings of the perceived intensity of suprathreshold odors are obtained. Almost uniformly, studies using all of these measures have shown a significant decline with age, typically beginning in the seventh or eighth decade of life. Age-related olfactory loss appears to develop gradually and is rarely complete, except in extreme old age. Nonetheless, it is often of sufficient magnitude to render older people vulnerable to chemical hazards such as gas leaks and to greatly diminish olfactory food flavor perception, reducing food enjoyment.

Despite extensive documentation of age-related decline in average olfactory sensitivity, debate continues regarding the uniformity of that decline, both across different odors and across individuals. Although there is little variation in the degree of loss reported for different odors, most studies have simply contrasted the performance of an elderly group with that of a group of young adults. As a result, possible differences in the onset or rate of decline in sensitivity to specific odors have not been examined. In addition, at the individual level, extreme differences among elderly subjects in olfactory abilities have frequently been noted, with some older individuals performing as well as the average young person. However, specific genetic, medical, and environmental factors that underlie this variation have not been identified, and some have questioned the reliability of apparent overlaps between the young and elderly in measures of olfactory ability.

Gustation

Taste receptors (responding to sweet, salty, sour, bitter, and umami, or "savory," stimuli) are also subject to damage from the chemicals they are designed to detect, as well as from viral, bacterial, and fungal species that often find a home in the oral cavity. However, taste receptors are not expressed on neurons, but on modified epithelial cells that turn over rapidly (approximately every two weeks). Moreover, these receptors are scattered over a large portion of the tongue dorsum, as well as being found on the soft palate, esophagus, pharynx, and epiglottis, and their responses are transmitted to the brain by multiple branches of three cranial nerves. These characteristics protect the taste system against extensive damage.

Actually, taste appears to be relatively stable across the life span. The measures used to assess gustatory function are similar to those used in studies of olfactory function, although few studies have included an assessment of the ability to identify the qualities of suprathreshold taste stimuli, at least partly because taste quality confusions (sour-bitter, sour-salty, and salty-bitter) are common even in healthy young subjects. Some age-related declines in both taste threshold sensitivity and the perceived intensity of suprathreshold tastes are typically observed; however, these declines are quality-specific and, in the case of bitter, compound-specific, and they are not always observed in both threshold and suprathreshold measures within a quality. Most modern studies have found no age-related decline in sensitivity to sweet (as exemplified primarily by sucrose), and declines in sensitivity to salty, sour, and bitter tastes, at either threshold or suprathreshold levels, are modest relative to those observed in smell. Because umami has only recently gained wide acceptance as a fifth basic taste, only a few studies have examined how it is affected by age. There are also few large, life-span studies of taste, but as is the case in smell, average declines appear to be significant only in the seventh and eighth decades.

All of these findings, however, are based on the whole-mouth presentation of taste stimuli. Several studies suggest the elderly are particularly prone to spotty losses of function affecting circumscribed areas of the tongue. In general, this has little impact on the whole-mouth experience of taste because other areas appear to compensate, but it may render elderly individuals more vulnerable to taste dysfunctions that lead them to seek medical assistance, including both whole-mouth loss of taste and the development of chronic, phantom taste sensations.

Beverly J. Cowart

See also Flavor; Olfaction; Taste; Taste System Structure

Further Readings

Cowart, B. J. (2005). Taste, our body's gustatory gatekeeper. *Cerebrum, 7,* 7–22. Retrieved from http://www.dana.org/news/cerebrum/detail.aspx?id=788

Cowart, B. J., Yokomukai, Y., & Beauchamp, G. K. (1994). Bitter taste in aging: Compound-specific decline in sensitivity. *Physiology & Behavior, 56,* 1237–1241.

Duffy, V. B., Cain, W. S., & Ferris, A. M. (1999). Measurement of sensitivity to olfactory food flavor: Application in a study of aging and dentures. *Chemical Senses, 24,* 671–677.

Murphy, C., Cain, W. S., & Hegsted, D. M. (Eds.). (1989). *Nutrition and the chemical senses in aging: Recent advances and current research needs* (Vol. 561). New York: Annals of the New York Academy of Sciences.

Murphy, C., Schubert, C. R., Cruickshanks, K. J., Klein, B. E. K., Klein, R., & Nondahl, D. M. (2002). Prevalence of olfactory impairment in older adults. *Journal of the American Medical Association, 288,* 2307–2312.

AGEING AND HEARING

A common complaint of older adults is difficulty hearing and understanding speech, which can result in problems following conversations, using the telephone, and detecting and locating alerting sounds. Approximately one third of adults ages 65 to 74 and nearly one half of adults 75 and older have impaired hearing. When loss of hearing for higher frequency sounds is included, nearly all individuals older than 80 have significant hearing loss. Indeed, hearing loss is among the most common chronic conditions of ageing, ranking first among men and fourth among women, after arthritis, cardiovascular disease, and cataracts. In the next few decades, the number of individuals with hearing loss will substantially increase as the population ages, as this entry describes.

Hearing loss in older adults is known as *presbyacusis* and results from genetic influences and the combined effects of ageing, long-term exposure to noise, ototoxic drugs, diet, trauma, and otologic and neurologic diseases. Additional risk factors may include smoking, cardiovascular disease, high blood pressure (hypertension), diabetes, and hyperlipidemia (high cholesterol). Presbyacusis can be narrowly defined as hearing loss that increases with age and is caused by the ageing process, which likely has a genetic basis. Because environmental and disease effects on hearing accumulate over a lifetime, hearing loss caused specifically by ageing is difficult to isolate in older humans. Older animals raised

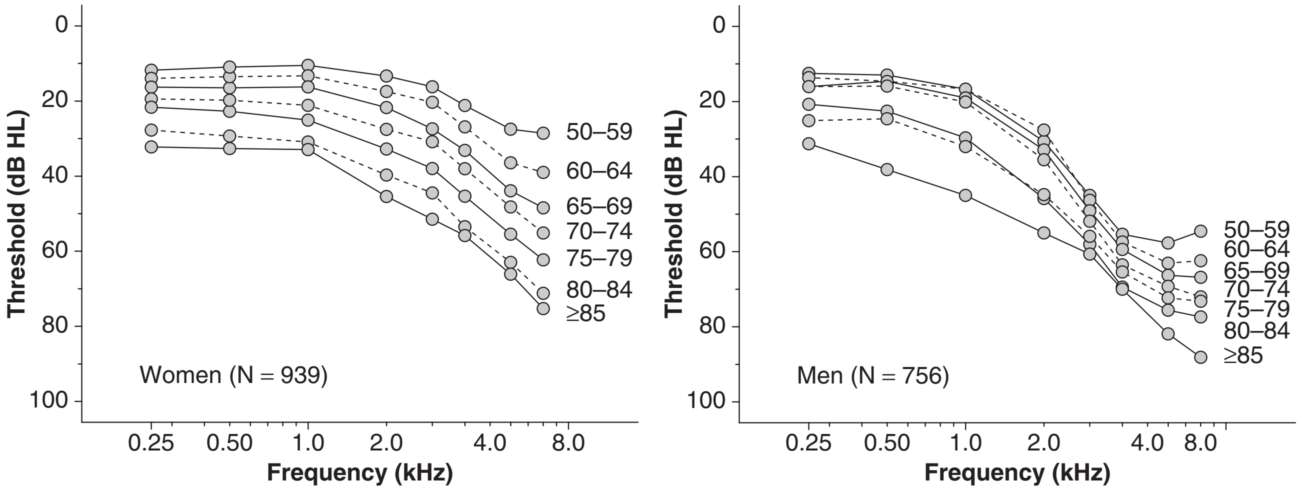


Figure 1 Thresholds for Pure Tones for Women and Men of Various Age Groups

Source: Medical University of South Carolina, Longitudinal Study of Age-Related Hearing Loss.

without exposure to noise or drugs nevertheless show unique age-related anatomic and physiologic changes in their auditory peripheral and central nervous systems (CNS). These include deterioration of tissues and blood supply in the cochlear lateral wall (inner ear), some degeneration of auditory neurons, and reductions in the resting potential that supplies power to the sensory (hair) cells.

Age-related hearing loss in the United States has been described by several community-based epidemiological and clinical studies. Figure 1 shows the systematic increase in hearing thresholds with increasing age in women and men, beginning with reduced ability to hear high-frequency and then mid-frequency sounds, which are important for understanding speech. Thresholds for women show the progression of a generally constant hearing loss at lower frequencies to a gradually sloping loss at higher frequencies; thresholds increase steadily with age more in higher than lower frequencies. Thresholds for men show similar patterns, but with more high-frequency hearing loss, which may relate to increased noise exposure.

Epidemiological surveys also suggest a strong genetic component in age-related hearing loss, which has been confirmed by results of twin studies and data from various strains of mice. Heritability coefficients indicate that as much as 55% of the variance in hearing thresholds of older persons is genetically determined, which is similar to heritability reported for hypertension and hyperlipidemia. The genetic component is reported to be stronger in women than men, which may be caused by a larger environmental component in men related to noise exposure. These results, along with studies of hearing in nonindustrialized societies, are consistent with the view that hearing loss in older adults results from environmental and disease factors superimposed on a genetically determined ageing process.

Along with declines in the ability to detect sounds (Figure 1), older adults have poorer abilities to detect small changes in intensity, frequency, and timing of sounds, which are important for using information carried by fluctuations in speech. The strongest evidence is from older adults whose thresholds are normal, which separates age-related effects from those resulting simply from poorer hearing. Differences between older and younger adults are larger at lower frequencies and decrease as frequency increases. This suggests a mechanism related to processing of timing information (e.g., detecting fast fluctuations or silent periods), which may degrade with age.

Many older adults complain of difficulty understanding speech, especially in noisy environments with many talkers. Declines in understanding of simple speech presented to one ear are largely explained by reduced speech audibility. Age-related differences not resulting simply from hearing loss are those for temporally distorted speech (such as rapid speech), speech presented binaurally with other sounds that are separated in space or fluctuating, or tasks requiring divided attention. Although behavioral measures of hearing and speech understanding are consistent with declines in processing timing information, it remains unclear whether these represent age-related changes in the auditory periphery, which provides impoverished signals to the CNS, or the combined effects of an ageing periphery, an ageing CNS, and cognitive factors.

Along with poorer communication abilities, untreated hearing loss in older adults can have negative effects on quality of life, contributing to social isolation, depression, and loss of self-esteem. Regrettably, no current medical treatment can prevent, delay, or reverse age-related hearing loss. The most widely used treatment is amplification provided by hearing aids and, for adults with severe-to-profound hearing loss, a cochlear implant, which is an electronic device implanted in the inner ear. Despite advances in hearing-aid technology, surveys report that more than 75% of individuals likely to benefit from a hearing aid do not own one (a gap of about 20 million people in the United States). Reasons are varied and include attitudes toward hearing loss and hearing aids, background noise and distortion, cost, adaptability, comfort, maintenance, tactile sensitivity, manual dexterity, and vision loss that may reduce benefits of speech reading.

Judy R. Dubno

See also Audition: Cognitive Influences; Audition: Disorders; Audition: Temporal Factors; Auditory System: Structure; Hearing Aids

Further Readings

Cruickshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, B. E., Klein, R., Mares-Perlman, J. A., & Nondahl, D. M. (1998). Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin: The epidemiology of hearing loss study. *American Journal of Epidemiology, 148,* 879–886.

Gates, G. A., Couropmitree, N. N., & Myers, R. H. (1999). Genetic associations in age-related hearing thresholds. *Archives of Otolaryngology-Head and Neck Surgery, 125,* 654–659.

Lee, F. S., Matthews, L. J., Dubno, J. R., & Mills, J. H. (2005). Longitudinal study of pure-tone thresholds in older persons. *Ear and Hearing, 26,* 1–11.

Schmiedt, R. A., Lang, H., Okamura, H. O., & Schulte, B. A. (2002). Effects of furosemide applied chronically to the round window: A model of metabolic presbycusis. *Journal of Neuroscience, 22,* 9643–9650.

Schuknecht, H. F., & Gacek, M. R. (1993). Cochlear pathology in presbycusis. *Annals of Otology, Rhinology & Laryngology, 102,* Supplement 158, 1–16.

AGEING AND TOUCH

Common wisdom indicates that sensory processes lose efficiency in old age. Glasses and hearing aids are a kind of standard aid for elderly people. Yet, in contrast to vision and hearing, the dramatic age-related deterioration of the sense of touch goes mostly unnoticed because there are no conditions such as reading newspapers or obtaining a driver license that would disclose an impairment of senses. As a result, the sense of touch and its vital role for coping with activities of daily living is widely underestimated. This entry discusses age effects on receptors and tactile perception, haptic perception, and treatability of age effects on touch.

Age Effects on Receptors

Touch perception is built on sensing through receptors, transmission through neurons at various stations along the sensory pathway, and central processing by the brain. As aging affects all components, multiple factors are involved in the age-related decline of the sense of touch.

Skin conformance undergoes alterations in old age. There is a loss of dermal receptors, which provide the senses of touch, pressure, and vibration. The morphology of Meissner's and Pacinian corpuscles, which detect vibrations and changes in texture change in old age, whereas Merkel-neurite touch receptors, which detect sustained touch and pressure, appear to be less affected.

Age Effects on Transmission and Central Processing

Transmission of touch information is affected by slowing of nerve conduction with increasing age and reduction of action-potential amplitude. In contrast to eyes and ears, which are close to the brain, the distances of skin receptors from the brain vary greatly across the body. Accordingly, skin receptors located on the feet are particularly vulnerable to a slowing down of conduction, which can amount to as much as 20 or 30%. It is believed that slow conduction emerging during aging results in greater temporal dispersion of incoming afferent information, which harms central processing thus contributing to perceptual impairment. At a central level, there is now agreement that neuron loss is a rare event. However, at a subcellular level, cortical sensory processing is affected through manifold age-related alterations of ion channels and receptor composition. Imaging studies in aged humans using magnetic resonance imaging indicated that gray matter density is reduced as a function of age, often starting as early as about 40 years of age; however, the time course of aging effects varies considerably over cortical areas. Thus, a major question is, what are the functional and perceptual consequences of these age-related changes, and how do these changes cause degradation of touch perception, which is discussed in the next paragraph.

Age Effects on Tactile Perception

At a perceptual level, a simple measure of tactile perceptual performance is tactile acuity, which describes the ability to resolve fine spatial details. Usually, measures of tactile acuity are obtained by testing two-point discrimination, gap detection, or grating orientation. Whatever technique is used, tactile acuity is significantly reduced during aging. For example, tactile acuity declined much more vigorously on some body regions than on others. According to one study, deterioration of acuity in the great toe averaged about 400% between young and elderly subjects (aged 65 to 87 years) compared with an average decline of 130% observed on the fingertip, meaning that acuity threshold is, depending on the body part, severalfold greater in older subjects (see also Figure 1, where spatial discrimination thresholds are plotted as a function of age).

Against intuitive thinking, the relation between receptor density and tactile acuity is far from being one-dimensional. Although the density of Meissner's corpuscles in the index and ring fingers does not differ, the acuity does. On the other hand, although males have lower density of Meissner's corpuscles compared with females, their acuity is not different. The reason for the lack of a simple relation between receptor density and perceptual outcome is that central somatosensory representations are not a mere mirror of the afferent input signals, but extract additional information through active computation based on the interaction of inhibitory and excitatory processes. In respect to the processing of touch information, there is a controversy about the role of peripheral versus central aspects that determine perception.

Haptic Perception

Providing tactile acuity information is only one among many functions the sense of touch subserves. Haptic perception, the ability to identify common objects by exploration with hands and fingers without sight, depends on accessing cutaneous information. For example, if afferent sensory information of the fingers is suppressed by wearing a glove or by local anesthesia of the fingers, motor control of the hand is severely impaired. In a study, both acuity thresholds and manual dexterity were assessed using a pegboard test that revealed a strong correlation between threshold and manual dexterity. As a result, age-related impairment of touch processing has far-reaching consequences on haptic and motor performance and may translate into great difficulties in tasks requiring fine manipulations. In addition to a reduction in tactile acuity, elderly people undergo a significant impairment in haptic perception.

In most cases, however, age-related decline of the sense of touch develops slowly and as part of a normal, that is, nonpathological, aging process over decades. As a result, elderly progressively adapt to the loss of high-level tactile performance and learn to compensate by developing behavioral strategies such as relying more on visual control to overcome the decrement in the sense of touch.

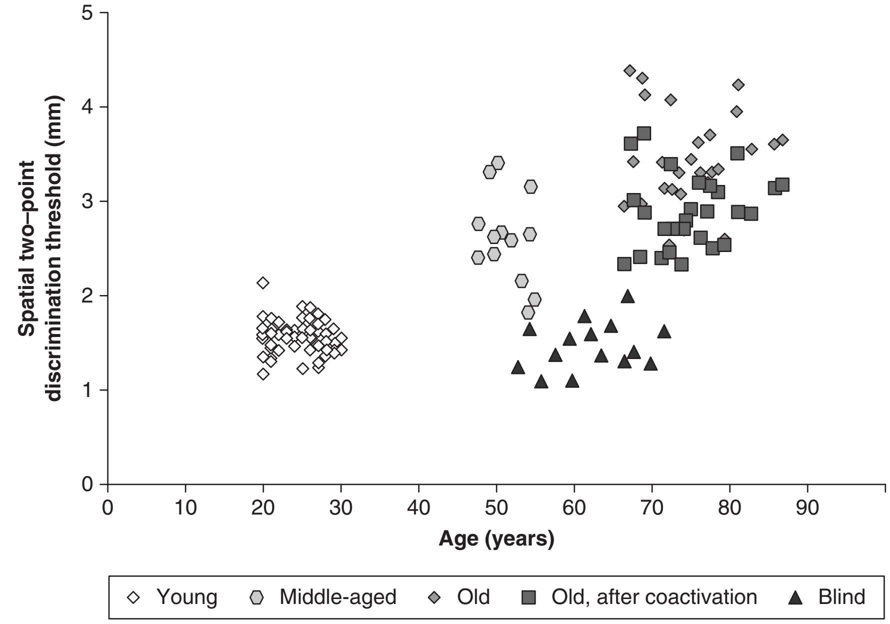


Figure 1 The Impact of Age, Training, and Sensory Stimulation on Spatial Discrimination Thresholds

Sources: Modified from Dinse, H. R., Kleibel, N., Kalisch, T., Ragert, P., Wilimzig, C., & Tegenthoff, M.(2006). Tactile coactivation resets age-related decline of human tactile discrimination. *Annals of Neurology, 60,* 88–94. Also from Heinisch, C., Kalisch, T., & Dinse, H. R. (2006). Tactile and learning abilities in early and late-blind subjects. In *Society for Neuroscience Abstracts 52.4/N2*. Washington DC: Society for Neuroscience.

Note: Each symbol represents an individual subject.

Treatability of Age Effects

Neuroplasticity is the ability of brains to adapt to new experiences and constraints of the environment during the entire life span. During the past few decades, such use-dependent plastic capacities of somatosensory systems have been intensively studied demonstrating that practice and training determine to a large extent perceptual abilities of an individual even at high age. This notion has been corroborated by recent findings demonstrating that the typical age-related decline in tactile performance is not inevitable, but that despite the accumulation of degenerative processes during aging, performance can be recovered by learning and training.

The impact of age, training, and sensory stimulation on tactile acuity is illustrated in Figure 1. Overall, discrimination performance is progressively impaired with increasing age. However, blind subjects who read Braille do not show the age-related changes, suggesting that maintained use of the fingers might play a role in preserving intact acuity. Furthermore, applying a brief period (3 hours) of tactile coactivation on the tip of the fingers, a procedure known to induce plastic changes in the somatosensory cortex, results in a significant amelioration of the age-related decline in elderly subjects, suggesting that age-related decline of tactile acuity is not irreversible. These results suggest that mechanoreceptor density may play a minor role in determining tactile acuity performance at old age but that acuity is rather controlled by cortical mechanisms. This view is also supported by recent imaging data that demonstrated a close correlation between tactile spatial resolution and the size of the area in the cortex devoted to processing of touch information. In conclusion, cortical processing, which is subject to alterations through practice and training, appears to be a crucial factor shaping the quality of touch-related performance even at high age.

Hubert R. Dinse, Martin Tegenthoff,
Christine Heinisch, and Tobias Kalisch

See also Action and Vision; Ageing and Hearing; Ageing and Vision; Cutaneous Perception; Experience-Dependent Plasticity; Haptics; Neuropsychology of Perception; Perceptual Learning; Proprioception; Reaching and Grasping; Sensory Rehabilitation; Tactile Acuity

Further Readings

Dinse, H. R., Kleibel, N., Kalisch, T., Ragert, P., Wilimzig, C., & Tegenthoff, M. (2006). Tactile coactivation resets age-related decline of human tactile discrimination. *Annals of Neurology, 60,* 88–94.

Kalisch, K., Ragert, P., Schwenkreis, P., Dinse, H. R., & Tegenthoff, M. (2008). Human age-related decline in tactile perception is accompanied by enlargement of the hand representation in somatosensory cortex. *Cerebral Cortex*. DOI: 10.1093/cercor/bhn190

Stevens, J. C., & Choo, K. K. (1996). Spatial acuity of the body surface over the life span. *Somatosensory and Motor Research, 13,* 153–166.

Tremblay, F., Wong, K., Sanderson, R., & Coté, L. (2003). Tactile spatial acuity in elderly persons: Assessment with grating domes and relationship with manual dexterity. *Somatosensory and Motor Research, 20,* 127–132.

AGEING AND VISION

Age-related changes in vision can profoundly affect the quality of life for seniors, decreasing independence and mobility and increasing physical injuries, such as hip fractures. But what aspects of vision change with age? And what produces those changes? Despite the critical importance of vision for functioning in everyday life, much still remains to be learned about the effects of healthy ageing on vision. This entry provides a brief overview of some of the most important aspects of ageing and vision researchers have learned.

Age-Related Effects

As people age, their vision is affected by changes both to their eyes and their brains. The average pupil size decreases with age, so that less light enters the eye, and this reduction in light is compounded by changes in the structure of the lens. Through the normal ageing process, even in the absence of cataracts (clouding of the lens), the lens becomes less transparent, blocking additional light from entering the eye, and changing the color of the light that does get through. Changes to the lens and intraocular media also increase light scatter, reducing the contrast of the retinal image. These structural changes to the lens also make it more difficult to focus on close objects as we age; this change in accommodative ability—the ability to change the optical power of the eye to focus on objects at different distances—is known as presbyopia, and is a normal part of the ageing process. The inability to focus on close objects without wearing glasses typically occurs between about 40 and 50 years of age, and the "near point" (the closest point on which one can focus) continues to move further away throughout our lifetimes. As young adults, people typically can focus clearly on objects about 10 centimeters away; by the time they reach 60 years of age, the near point has moved away to about 1 meter. Fortunately, by wearing appropriate optical corrections (glasses), seniors can overcome many of the difficulties associated with presbyopia.

These optical changes adversely affect vision, but research has shown that the most profound age-related effects on vision stem from changes in neural, rather than optical, processing. Blurring of the optics would lead to an inability to see fine details, resulting in a decrease in visual acuity. But ageing also decreases our contrast sensitivity—defined as the inverse of the lowest contrast needed to see different sized targets—even for mid-sized objects, although the most pronounced effects are seen for objects with more fine details. For example, the general shape and features of a face would remain visible throughout our lives, but it becomes more difficult to discriminate individual eyelashes around the eye, or fine wrinkles on a face. Based on behavioral techniques that can tease optical and neural factors apart, and based on studies in which stimuli are projected directly onto the retina, bypassing the optics, the age-related reduction in contrast sensitivity is largely the result of changes in neural processing, rather than of changes in the optics of the eye. Regardless of what causes a change in our acuity and contrast sensitivity, the net result is that the world appears a bit more blurry and washed out in the elderly.

Ageing affects many stages of neural processing of visual information, beginning with a reduction of sensitivity in cone photoreceptors (the cells that enable us to discriminate color) in the retina. Interestingly, however, although the relative sensitivity of the cones changes continuously from adolescence through our senior years, color perception remains relatively constant, suggesting that the neural systems encoding color engage compensatory mechanisms, although the precise nature of such mechanisms remains unclear. Recent physiological studies have also suggested that the intrinsic level of noise in visual neurons is increased and the selectivity of visual neurons to visual attributes such as direction of motion and orientation decreases in senescent monkeys. Such changes may be because, throughout the brain, there are changes in the quantity and function of certain neurotransmitters. For example, the efficacy and production of the brain's primary inhibitory neurotransmitter (gamma-aminobutyric acid, or GABA) decreases with age, at least in some cortical areas.

Implications for Task Performance

As might be expected from the neural changes in the selectivity of motion mechanisms, older adults are impaired on many motion perception tasks. For example, older observers can detect only about a 10% difference in speed of a cluster of moving dots or other targets, whereas younger observers can detect speed differences half that size (5%). Similar age-related effects are seen for direction discrimination, although there is some indication that these basic motion deficits mostly affect "senior-seniors" (observers older than 70 years of age), rather than for "junior-seniors" (60–69). In some unusual circumstances, older observers can even outperform younger observers. For example, seniors are better able and faster than young observers to detect the direction of motion for large high-contrast targets. Although this result seems counterintuitive, it may be a by-product of reduced intra-cortical inhibition (e.g., GABA levels) in the visual cortex. Inhibition in the brain typically highlights objects on backgrounds, and "tunes out" large, uniform patterns, resulting in decreased sensitivity to motion for large, high-contrast targets in young observers. If inhibition declines with age, large, high-contrast targets would not be suppressed as much in seniors, and, consequently, detection of motion of such targets should improve with age. In general, however, seniors consistently have difficulties in perceiving more complex forms of motion such as shape from motion, optic flow (the way objects and scenes move across the retina as they or the observer move), biological motion (e.g., the ability to discern the presence of a person walking when viewing only a sparse array of dots on the locations of a person's joints), and motion defined across different visual attributes. Evidence also suggests that older women are more impaired in a range of motion tasks than are older men, although the extent of this sex difference and the mechanism underlying it are unknown.

The effect of stimulus and task complexity on ageing is not limited to motion perception. In general, age-related deficits appear to grow with increasing stimulus or task complexity. For example, ageing seems to impair our ability to integrate information across space, even when our perception of local elements remains intact. A good example of this is seen in the perception of oriented contours. Older observers have little if any difficulty in discriminating localized, oriented targets; indeed, in humans, the behaviorally estimated neural selectivity of orientation mechanisms seems unaffected by age (a result that stands in contrast to the finding from single-cell physiology in monkeys suggesting that ageing may decrease orientation selectivity, at least in anesthetized animals). However, integration of local orientations into larger, more spatially distributed contours appears to be impaired. For example, young adults have no difficulty grouping several small oriented lines together to form a curve, but seniors appear not to be able to integrate those local orientations together to form a global shape as well as younger observers can. Other aspects of visual integration, including the perception of symmetry, are similarly impaired with age, and older observers are also impaired on perceptual tasks that require observers to divide their visual attention across space or to detect a target in a cluttered environment (evidenced in the lab by relatively poor performance by seniors on visual search tasks in which observers have to find a specific object hidden among a number of other objects).

Why might age have a larger impact on visual processing for complex rather than for simple stimuli and tasks? One possibility is that compensatory mechanisms are active in older brains, but that there are limits to the extent of that compensation. Recent advances in neuroimaging techniques shed some light on this hypothesis. Researchers now know that even when younger and older observers perform identically on a simple visual discrimination task, the neural systems engaged in younger and older brains may be quite distinct. For example, in one study, whereas the visual cortex was of primary importance for a visual discrimination task in younger observers, older observers performing the same task appeared to recruit parts of the hippocampus (normally associated with memory) and prefrontal cortex (normally associated with attention).

Overcoming Limitations

Fortunately, the visual system remains tremendously plastic throughout our lifetimes, so that seniors can continue to learn new visual tasks and can learn to improve performance on more familiar tasks. For example, as mentioned earlier, seniors typically are less able to divide their attention across space than younger observers (e.g., in a multitasking, divided attention task, in which observers must name a central letter while detecting a spot in the periphery). With about a week's worth of training, however, seniors can learn to divide attention as well as younger observers do. Effects of training can be seen for lower-level aspects of vision as well, including the discrimination of moving stimuli. In some cases, such learning can last for months, even with no intervening training. Important questions for future research include determining ways in which learning can be optimized, what the effects of learning on the brain are, and how age-related changes in vision (and improvements based on learning) affect seniors' ability to function in the everyday world.

So, although the visual world of the senior may be more of a "blooming, buzzing confusion" than is that of the young adult, older brains can learn new tricks and can compensate in a variety of visually challenging situations. That, combined with glasses that can correct many optical problems, enables seniors to overcome many of the limitations in vision that accompany healthy ageing.

Allison B. Sekuler

See also Ageing and Chemical Senses; Ageing and Hearing; Ageing and Touch; Color Perception; Depth Perception in Pictures/Film; Motion Perception; Vision; Visual Acuity

Further Readings

Andersen, G. J., & Atchley, P. (1995). Age-related differences in the detection of three-dimensional surfaces from optic flow. *Psychology and Aging, 10,* 650–658.

Betts, L. R., Taylor, C. P., Sekuler, A. B., & Bennett, P. J. (2005). Aging reduces center-surround antagonism in visual motion processing. *Neuron, 45,* 361–366.

Enoch, J. M, Werner, J. S., Haegerstrom-Portnoy, G., Lakshminarayanan, V., & Rynders, M. (1999). Forever young: Visual functions not affected or minimally affected by aging: A review. *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 54,* B336–351.

Faubert, J. (2002). Visual perception and aging. *Canadian Journal of Experimental Psychology, 56,* 164–176.

McIntosh, A. R., Sekuler, A. B., Penpeci, C. Rajah, M. N., Grady, C. L., Sekuler, R., & Bennett, P. J. (1999). Recruitment of unique neural systems to support visual memory in normal aging. *Current Biology, 9,* 1275–1278.

Norman, J. F., Ross, H. E., Hawkes, L. M., & Long, J. R. (2003). Aging and the perception of speed. *Perception, 32,* 85–96.

Owsley, C., Sekuler, R., & Siemsen, D. (1983). Contrast sensitivity throughout adulthood. *Vision Research, 23,* 689–699.

AGNOSIA: AUDITORY

Auditory agnosia is the loss of auditory recognition in the absence of deafness. It is often associated with lesions of the auditory cortex, but the definition does not require a cortical lesion. It also is possible to have auditory agnosia resulting from disorders of the auditory nerve, brainstem, and midbrain. Acquired agnosia following brain insult is rare because it often, but not always, follows damage to both sides of the brain, which is usually fatal. The assessment of auditory agnosia requires the assessment of both discrimination and recognition. Problems with discrimination occur in *apperceptive agnosia* (caused by loss of perception), when there will also be deficits in recognition. In *associative agnosia,* discrimination (and therefore perception) is intact, but there will be a deficit in recognition caused by a deficit in the attribution of semantic labels to sounds. The two general types of auditory agnosia are therefore disorders occurring at different levels in auditory object analysis. Apperceptive and associative forms of agnosia for word recognition (pure word deafness) and voice recognition (phonagnosia) are both described in this entry. Agnosia for music (amusia) can be characterized as an apperceptive agnosia. Environmental sound agnosia has also been described.

Causes

Auditory agnosia is most commonly caused by strokes on both sides affecting the upper temporal lobe. It can also occur after temporal lobe damage resulting from herpes simplex encephalitis. Many cases of auditory agnosia reported after cortical lesions do not actually meet a strict definition of agnosia based on the presence of normal hearing. This may reflect the evolution from cortical deafness to auditory agnosia that is observed in a number of cases. In addition, many patients with auditory agnosia caused by stroke are from an elderly population in which deafness resulting from abnormal cochlea function is common.

Lifelong (congenital) forms of agnosia have also been described. In particular, the disorder commonly known as tone deafness or tune deafness can be characterized as a congenital agnosia called congenital amusia.

Overlap Between Apperceptive Agnosias

In apperceptive forms of agnosia, there is often an overlap between the domains of perception that are affected. Deficits in the recognition of single words (word deafness), environmental sounds (environmental-sound agnosia), and music (amusia) commonly coexist, especially in patients with agnosia resulting from damage to the upper temporal lobe on both sides. This reflects overlap in the aspects of auditory perception required for the analysis of these different sounds, where the profile of specific deficits depends on the particular aspect of auditory analysis that is affected by the cortical lesions. The degree of overlap between domains affected by agnosia is likely to be underestimated in view of the practical difficulty of carrying out a comprehensive and rigorous assessment of all of these, and that it is unusual for patients to undergo a systematic assessment of complex sound analysis.

Some studies of patients with apperceptive forms of agnosia have examined temporal aspects of complex sound perception, especially in patients with word deafness, using approaches that include click counting and discrimination of tone sequences or rhythms. Many of the reported patients with word deafness in which such temporal deficits have been described also suffered deficits in other domains, as would be predicted on the basis of an apperceptive agnosia caused by a deficit in temporal analysis. Deficits in examining the frequency structure (spectrum) of sound have not been systematically explored in agnosia, but might be relevant to deficits in musical perception, in particular, which is less robust than speech is to degraded spectral structure.

The damage associated with auditory agnosia commonly affects the primary and secondary auditory cortex in Heschl's gyrus in the uppermost part of the temporal lobes and the association cortex behind Heschl's gyrus in a region called the planum temporale. Studies of normal subjects using functional imaging suggest that different aspects of time analysis and spectral analysis occur in different regions of the auditory cortex, and different distributions of damage in different patients can cause different types of deficits in complex sound analysis and therefore different differential effects on speech, music, and environmental sound perception. Attentional effects in association with such damage are also common and will contribute to the clinical defect observed.

Agnosia for Words and Voices

The distinction between acquired apperceptive and associative agnosia has been emphasized in the word-deafness literature. Associative forms of word deafness, where word discrimination is intact but word recognition impaired, can occur with damage to the left temporal lobe, although left-side lesions have also been reported to produce an apperceptive agnosia. A similar distinction has been made in the case of phonagnosia: agnosia for voices. Reports of symptomatic deficits in voice recognition are rare in the case literature, but in series of asymptomatic patients with damage to the right or left temporal and parietal lobes, apperceptive phonagnosia has been reported resulting from right or left temporal lobe damage and associative phonagnosia resulting from right parietal lobe damage.

Agnosia for Music

Musical agnosia (amusia) has been studied in a number of patients following unilateral or bilateral temporal lobe stroke. It can be argued that musical deficits resulting from stroke can be considered relative to the components of music in the pitch or rhythm domain. In the pitch domain, for example, deficits can occur in the perception of the pitch of individual notes, simple pitch patterns such as pitch-interval direction between pairs of notes, contour (the patterns of "ups" and "downs" in a melody), and tonal structure. Neuropsychological models suggest that these are successive levels in a hierarchy of pitch and pitch-pattern analysis, and normal functional imaging suggests that the substrate for such analysis becomes increasingly distributed (involves areas increasingly distant from the primary and secondary auditory cortex) when moving up the perceptual hierarchy. The disorder commonly coexists with a deficit in the perception of prosody in speech at a similar level of temporal structure to the critical level in music: prosody involves changes in pitch, stress, and rhythm within a temporal "window" that is also at the level of hundreds of milliseconds. The disorder is best characterized as an apperceptive agnosia before the level of deficits in tonal analysis, when learned associations with pitch patterns are made and a form of associative deficit might be argued.

In the pitch domain, there is a tendency for deficits at higher levels in the pitch hierarchy to be caused by lesions that are more remote from the primary and secondary auditory cortex. Right-hemisphere lesions causing deficits are more commonly reported and therefore emphasized in the literature but left-hemisphere lesions *can* produce deficits in mechanisms for the analysis of pitch and pitch pattern. In part, this might reflect an ascertainment bias. Cases of acquired amusia resulting from left-hemisphere damage may be associated with aphasia, making behavioral assessment more difficult.

There has been recent interest in lifelong disorders of musical perception, known as tone deafness, tune deafness, or congenital amusia. Lifelong disorders of singing have been recognized for more than a century, and recent work using a schedule called the Montreal Battery for the Assessment of Amusia (MBEA) demonstrates that the deficit is commonly associated with a deficit in musical perception that can be demonstrated by impaired discrimination of novel melodies. The disorder can therefore be characterized as a congenital form of apperceptive agnosia and is particularly associated with deficits in the pitch domain. Rhythm deficits can also occur, but only if these occur in the context of a melody. Behavioral studies have demonstrated deficits in pitch discrimination and pitch-direction analysis, suggesting a fundamental deficit in the analysis of pitch pattern. Based on lesion data from a number of studies and consideration of the normal cortical mechanisms for pitch-pattern analysis demonstrated by functional imaging, a cortical basis for the disorder would be predicted. Although abnormalities of brain structure are not seen at the individual level, group studies of brain structure using magnetic resonance imaging provide evidence for cortical disorder abnormalities that particularly affect mechanisms in the right temporal lobe and frontal lobe for the analysis of pitch pattern and pitch working memory for pitch: the mechanism by which we hold notes in mind when we listen to musical phrases.

Timothy D. Griffiths

See also Agnosia: Tactile; Agnosia: Visual; Music Cognition and Perception; Neuropsychology of Perception

Further Readings

Griffiths, T. D., Bamiou, D. E., & Warren, J. D. (2009). Disorders of the auditory brain. In A. Rees & A. R. Palmer (Eds.), *Oxford handbook of auditory science: The auditory brain.* Oxford, UK: Oxford University Press.

Griffiths, T. D., Kumar, S., & von Kriegstein, K. (2009). Auditory object analysis. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences.* Cambridge: MIT Press.

AGNOSIA: TACTILE

Everyday object recognition involves contributions from all of our senses. We recognize objects not only by their shape and color, but also by their sounds, textures, and temperatures. Our brains are organized to integrate information across senses, but each sense has its own separate neural object recognition system. To understand how the brain achieves multisensory object recognition, researchers examine how brain damage leads to the selective failure of recognition by an individual sensory system. This entry focuses on how impairments of tactile object recognition, or tactile agnosia, contribute to our understanding of normal object recognition. The entry reviews issues regarding different types of tactile agnosia, brain regions involved in tactile agnosia, sensory and motor contributions to tactile agnosia, and the existence of tactile agnosia separate from visuospatial disorders.

Agnosia is defined as the impaired recognition of familiar objects, faces, or sounds, despite adequate perceptual and intellectual capabilities. *Tactile agnosia,* or *somatosensory agnosia,* refers to a deficit in recognizing common objects by touch following brain damage, despite sufficiently intact tactile sensation, memory, and general intellectual function. Tactile agnosia is distinct from *tactile aphasia,* an inability to verbally label a felt object, or tactile inexperience in which the patient is not familiar with the object.

Tactile agnosia is modality specific: Patients who have difficulty recognizing objects by touch can still recognize objects by vision, audition, taste, or smell. Patients describe their condition as feeling like their hand is "numb" or "stiff," that the feeling is not distinct or strong enough to identify the objects by touch. Patients with visual agnosia have similar complaints saying their glasses appear "dirty" or "don't work."

Patients with tactile agnosia usually have unilateral lesions (i.e., brain damage on only one side of the brain). The hand that is contralateral to the lesion (on the opposite side of the lesion) is the hand that is "agnostic." A patient with a left hemisphere lesion will have difficulty recognizing objects with his or her right hand. The unilateral nature of tactile agnosia provides an advantage to investigating tactile agnosia: patients provide their own control comparisons because patients' ability to recognize felt objects with their non-agnostic, ipsilateral hand (i.e., on the same side of the lesion) can be compared directly with their inability to recognize the same objects and perform the same tasks with their contralateral agnostic hand. Thus, difficulties recognizing objects with their agnostic hand cannot be attributed to labeling problems or a lack of knowledge about the objects.

Types of Tactile Agnosia

To identify different types of tactile agnosia, researchers consider stages of processing for tactile object recognition. Tactile object recognition is assessed by asking patients to feel objects that they cannot see or hear, and identify them. When feeling objects, patients usually explore the object with one hand; this active touch is called *haptic perception.* Patients may also recognize patterns passively, without hand movement. Identification of felt objects includes verbally providing a name or label for an object (e.g., cup), pointing to a visually presented object or picture, and matching the felt object with another object.

In theories of tactile object recognition, perceptual input information is transformed through a series of processing stages before object identification. The first stage of processing involves the extraction and processing of the properties and features of the felt object. For example, to recognize a dime, one must first apprehend its small size, cool temperature, round shape, and rough edges. Exploratory procedures, or hand movements, integrate input from cutaneous, kinesthetic, and proprioceptive receptors to extract information about the object's texture, hardness, shape, size, weight, part relations, and temperature. Next, these properties must be combined or integrated into a common modality-specific representation or object description. The integrated tactile object description is then used to access information about object meaning and function (i.e., semantics). These descriptions contain semantic information from multiple senses that permit object recognition.

Object recognition deficits can be associated with earlier or later stages. If the property integration stage (tactile attributes cannot be combined with each other) is disrupted, then patients will be able to perceive the object's features but will not be able to draw the objects or describe how the various features relate to each other. If patients cannot associate the current perception of the object with its stored meaning (i.e., the access to semantic representation stage is disrupted), then patients can describe the object and draw it, but not be able to identify it. This division corresponds to the two major types of deficits revealed in case studies of both visual and tactile agnosia: apperceptive and associative.

Apperceptive tactile agnosia refers to tactile object recognition deficits in higher-level tactile perception. Patients with apperceptive tactile agnosia have relatively intact basic tactile perception in that that their abilities to discriminate two points on their hands, feel pressure, light touch, and finger movement are within the range of non-brain-damaged individuals of the same age. They can distinguish between different tactile textures, compliant surfaces, temperatures, and sizes. However, they have difficulty forming a perceptual representation, or percept, that integrates the individual tactile object properties, especially object shape and part relations. When feeling a combination lock, a patient with apperceptive tactile agnosia may describe its hook on top, its rounded shape, its moving dial, and its raised number ticks on the dial, but will be unable to put the parts together to identify the object. Further, if the patient is asked to draw the felt object, the drawing tends to have the wrong proportions or numbers of the object's parts (e.g., a cassette tape with twelve instead of two holes) or to only capture a vague outline of the object that lacks interior detail (e.g., no ticks on the dial of a combination lock).

Associative tactile agnosia refers to deficits that occur when patients cannot use the integrated tactile percept to access their knowledge of the object. In contrast to apperceptive tactile agnosia, patients with associative tactile agnosia can verbally describe the material and shape of felt objects. They can make drawings of tactually presented objects as well as match two- and three-dimensional shapes. Despite these capabilities, they cannot recognize the objects, determine whether two felt objects are from the same category (e.g., they are both necklaces), or both perform the same function (e.g., a cork screw and a church key both open bottles). As a result, their inability to recognize objects by touch is interpreted as disrupted connections between incoming perceptual information and stored object knowledge or semantic representations. Although apperceptive and associative agnosias provide an important distinction, current case studies refine these distinctions by studying the precise stage of processing at which the specific deficit occurs.

Cortical Bases for Tactile Agnosia

Case studies implicate the parietal lobe as the critical neural region for tactile agnosia. Patients with tactile agnosia often have lesions in the second somatosensory cortex (SII) and inferior parietal regions (Brodmann areas 39 and 40). The primary somatosensory cortex (SI) processes individual attributes such as values texture, hardness, vibration, and temperature on a particular body part, but these regions appear to integrate object attributes. Recent neuroimaging studies of non-brain-damaged individuals have confirmed these parietal regions as important for the initial, somatosensory-specific stages of tactile object recognition; later stages of tactile object recognition processing appear to follow the ventral temporal pathway used by the visual object recognition system. Thus, the initial stages of tactile object recognition are specific to tactile processing and the later stages are shared at least by the visual system, suggesting that information from individual object recognition systems may converge in the ventral or "what" pathway of the temporal lobe.

Sensory and Motoric Contributions

Similar to the study of visual agnosia, researchers debate whether tactile agnosia can exist without sensory impairment, and if so, how sensory and motor deficits influence the recognition process. The current consensus is that tactile object recognition can be disrupted without clinically demonstrated somatosensory loss. Supporting this view, some patients with tactile agnosia have shown comparable performance with non-brain-damaged individuals on basic and intermediate somatosensory functions despite severe tactile object recognition deficits. Tests of basic somatosensory functions include light touch (with a wisp of cotton), vibratory sensation (on a joint), proprioception (determine the direction the experimenter moved a joint), superficial pain (pin prick), temperature (cold/warm object applied to skin), and two-point discrimination (distance that two touches can be distinguished on the skin). Intermediate somatosensory functions include determining an object's texture, size, weight, hardness, and simple shape judgments (long or short). Thus, tactile agnosia is a higher-level processing problem that cannot be attributed to perceptual problems.

To examine the extent to which somatosensory and motoric inputs influence the tactile object recognition process, researchers have examined individuals with sensory and motor deficits. Patients who have damage to their hands or peripheral nervous system and patients with lesions in SI do not always show tactile object recognition deficits in their contralesional hand. Likewise, patients with hand paralysis do not necessarily have significant deficits in tactile object recognition. Touch differs from other senses because it interacts directly with objects themselves and most people perform exploratory procedures to actively investigate an object. For example, to determine an object's texture, they tend to rub the object; to determine its hardness, they squeeze it; and to determine an object's shape, they grasp it or run their fingers around the edge of it. Because the hand uses both tactile inputs and hand movements to extract somatosensory information, these patients probably are using the motions of object parts or hand movement cues to extract relevant object information that they cannot obtain through purely tactile perception. In sum, there are sensory and hand movement contributions to tactile agnosia, but they cannot explain the overall deficit.

Spatial Contributions

Researchers have also questioned whether tactile agnosia exists independently of a more general, supramodal (not specific to a particular modality) spatial impairment. This question arose from the result of a large study that compared the performance of brain-damaged World War II veterans with non-brain-damaged patients on a variety of tests of tactile and spatial function. Tactile tests include two- and three-dimensional pattern recognition. Spatial tests included a map-following task in which people were given a map and room with markers on the floor; they were asked to walk the path indicated by the map. Results showed that brain-damaged patients who were impaired on the tactile tests were also impaired on the spatial tests, suggesting that the two abilities had common neural sources. However, recent research has documented that tactile agnosia can exist without concurrent spatial deficits. A patient with documented tactile agnosia was assessed on visual spatial tests designed to assess the same spatial integration process assumed to be taxed by tactile object recognition, including the same map test used in the earlier study. Despite a severe tactile object recognition deficit, the patient performed at above-average levels on all spatial tests, both tactile and visual. This case study indicated that tactile object recognition problems may be specific to the touch and are not caused by general spatial deficits such as the inability to manipulate complex spatial images or to imagine how object parts fit together. Consistent with this conclusion, neuroimaging research has confirmed that the parietal regions involved in tactile object recognition are different from those involved with spatial localization. This functional dissociation in the somatosensory system of object recognition ("what") and spatial processing ("where") is similar to what is found in the visual system. The conflict between the results of these two studies may be resolved by the fact that the locations and extent of the lesions were unknown in the older study. The lesions of those patients may have involved both functional regions of the parietal lobe.

Significance

The study of tactile agnosia reveals that our brains have neural systems specialized for tactile object recognition that are independent of visual and auditory object recognition systems. Nonetheless, there are commonalities in the tactile and visual stages of object recognition at both functional and neural levels. These similarities provide a basis for the interaction of the two systems that permits our normal, multisensory recognition of common objects.

Catherine L. Reed

See also Agnosia: Auditory; Agnosia: Visual; Cutaneous Perception; Cutaneous Perception: Physiology; Haptics; Neuropsychology of Perception; Object Perception; Recognition

Further Readings

Bauer, R. M., & Rubens, A. B. (1985). Agnosia. In K. M. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (2nd ed., pp. 187–241). New York: Oxford University Press.

Caselli, R. J. (1991). Rediscovering tactile agnosia. *Mayo Clinic Proceedings, 66,* 357–364.

Endo, K., Miyasaka, M., Makishita, H., Yangisawa, N., & Sugishita, M. (1992). Tactile agnosia and tactile aphasia: Symptomatological and anatomical differences. *Cortex, 28,* 445–469.

Nakamura, J., Endo, K., Sumida, T., & Hasegawa, T. (1998). Bilateral tactile agnosia: A case report. *Cortex, 34,* 375–388.

Platz, T. (1996). Tactile agnosia: Casuistic evidence and theoretical remarks on modality-specific meaning representations and sensorimotor integration. *Brain, 119,* 1565–1574.

Reed, C. L., & Caselli, R. J. (1994). The nature of tactile agnosia: A case study. *Neuropsychologia, 32,* 527–539.

Reed, C. L., Caselli, R. J., & Farah, M. J. (1996). Tactile agnosia: Underlying impairment and implications for normal tactile object recognition. *Brain, 119,* 875–888.

Semmes, J. (1965). A non-tactual factor in astereonosis. *Neuropsychologia, 3,* 295–315.

Srinivas, K., & Ogas, J. (1999). Disorders of somesthetic recognition: A theoretical review. *Neurocase, 5,* 83–93.

AGNOSIA: VISUAL

Visual agnosia refers to a class of disorders in the domain of visual recognition, in which the affected person cannot arrive at the meaning of some or all categories of visual stimuli. In such cases, knowledge of the shape, form, or other stored knowledge of the physical attributes of an object cannot be accessed, with the result that the stimulus is not successfully identified. The failure to recognize the object occurs under a whole range of conditions and independent of whether the stimulus is presented as a real three-dimensional object, as a black-and-white line drawing, or as a photograph. Dramatically, these perceptual failures occur despite the individual exhibiting normal or near-normal elementary visual functions such as acuity, brightness discrimination, and color vision, along with normal or near-normal semantic and memory functioning. These individuals also have intact alertness, intelligence, and language, thus setting aside questions about whether agnosia is simply a manifestation of reduced elementary visual function and intelligence.

Visual agnosia may result from a lesion (e.g., as a consequence of a stroke or tumor or other form of brain damage) sustained by an adult who possessed normal premorbid perceptual abilities, or it may be apparent in a child who sustained a brain lesion early in life, a disorder usually referred to as *developmental agnosia.* Visual agnosia may also occur in the context of a progressive deterioration of perceptual skills, as in the visually selective progressive posterior cortical atrophy, and finally, it may be evident in individuals who are apparently impaired at recognition (predominantly of faces) from birth in the absence of any obvious neurological concomitant (see Table 1).

Not surprisingly, patients may fail to recognize objects if they have a deficit in semantics (knowledge or meaning of objects and their properties such as what an object might be used for). Interestingly, however, and the focus of this short review, agnosic patients typically retain the long-term representation of the object but still fail to recognize the stimulus presented. Thus, several investigators have reported data from visual imagery tasks that demonstrate that stored visual knowledge may be at least relatively preserved in patients with severe visual agnosia and that, despite this, the patients do not recognize the presented object; for example, patients may be able to visualize in their mind's eye a particular object but, presented with the same object, may fail to identify it. Consistent with this, these patients usually demonstrate normal recognition of objects through modalities other than vision (touch, audition, verbal definition, or description of its function), further indicating that the deficit is not simply a difficulty in retrieving names or in accessing the necessary semantic/knowledge information. In short, the problem appears to be one of accessing meaning from visual input.

Visual agnosia can be general, affecting the recognition of all visual stimuli, or it can be more specific: for example, some agnosias are (relatively selective) for objects, for faces or *prosopagnosia,* for words or *pure alexia,* for body parts, for colors, and for environmental scenes, including landmarks. Finally, the disorder of simultanagnosia—an

Table 1 Classification of Types of Visual Agnosia, Their Underlying Neuropathology, and Clinical Manifestation

	Neuropathology	*Clinical Manifestation*
Apperceptive agnosia	Stroke, anoxia, carbon monoxide poisoning affecting occipital, parietal, or posterior temporal regions bilaterally	Unable to copy, match, or identify visual stimuli
Associative agnosia	Usually bilateral infarction of the posterior cerebral arteries but unilateral occipitotemporal damage may suffice	Able to copy and match stimuli, may even be able to provide verbal description of aspects but still not recognize the object
Integrative agnosia	Extensive extrastriate damage bilaterally or to the right hemisphere	Fails to piece the components of an object together so oversegments or perceives in segmental fashion
Prosopagnosia	Acquired form: bilateral infero-mesial visual association cortices (lingual and fusiform gyri) and subjacent white matter	Failure to recognize faces
Agnosia for words	Left occipitotemporal cortex	"Pure" alexia—fails to read words normally given normal language and sensory visual function
Agnosia for scenes	Bilateral or right posterior artery infarction involving the fusiform and lingual gyri, extending to the parahippocampal gyrus	Failure to recognize landmarks or known scenes
Developmental agnosia	No obvious neural concomitant on clinical scanning	Difficulty in acquiring mastery over word or face recognition (developmental dyslexia or developmental/congenital prosopagnosia)
Posterior cortical agnosia	Focal right temporal or occipital lobe atrophy or both	Progressive decline in complex visual functions and recognition

inability to "see" more than one object at a time—is also regarded as a type of visual agnosia. Some cases evince fairly "pure" forms of one of these agnosias, but there are many reported cases in whom some subset of the specific forms can co-occur. Some of the more common subtypes of visual agnosia are described in this entry.

Visual Object Agnosia

The term *visual agnosia* was coined by Sigmund Freud to refer to the problems observed in some individuals with object recognition problems, but is classically divided into two broad categories: *apperceptive mindblindness* and *associative mindblindness*. A person with apperceptive agnosia is impaired at constructing a perceptual representation from the visual input and, consequently, is unable to copy, match, or identify a drawing based on the impoverished visual representation derived from the image. Apperceptive agnosia corresponds to the breakdown at the stage of visual processing at which the elementary features of the stimulus are processed and its structural description is achieved—a relatively early stage of the visual recognition system (see Table 1 for description of all agnosia types as well as neuropathologic basis and clinical manifestation). In contrast, a person with associative agnosia cannot use the well-specified perceptual representation to access stored knowledge of the object's functions and associations—such an individual is able to copy and match the object (or a drawing/rendition of it) even though unable to identify it (this disorder is often described as *perception stripped of meaning,* to borrow the phrase from Hans-Lukas Teuber). Whether the derived percept is truly normal in associative agnosia and simply cannot be used to access meaning or whether it is still not sufficiently well derived is a topic of much contention.

One reason that the visual agnosias are so intriguing a class of clinical phenomena is that they have important implications for current theories of high-level vision. In the service of elucidating normal perception, then, more fine-grained neuropsychological accounts as well as computational accounts have fractionated the classical distinction between apperceptive and associative agnosia further because of the growing understanding that visual object recognition comprises a number of distinct steps, in transforming the retinal input into representations that reflect invariant properties of objects in the real world. In the course of doing so, these authors have identified agnosias that arise at different stages of object processing. One such addition is *integrative agnosia.*

Patients with integrative agnosia appear to have available to them the basic features or elements in a display but are unable to integrate all aspects into a meaningful whole. These patients operate in a piecemeal manner to identify objects, sometimes over-segmenting the single input object into several different objects (e.g., identifying the handle of a fork as one object and the tines as a second). These individuals may perform normally on visual matching and copying tasks but perform poorly when required to bind visual elements in a spatially parallel fashion across a field containing multiple stimuli. The failure to integrate the disparate elements occurs equally with displays of two- and three-dimensional stimuli and with black-and-white and chromatic displays; in some cases, however, the presence of depth, color, and surface cues may be of some assistance to the patients in segmenting the display. These integration/segmentation problems are most clearly demonstrated when there are multiple items present, such that there is competition in assigning elements between shapes, as in displays with overlapping shapes or when boundary assignment is required. Interestingly, in some patients, the presence of local information is so captivating that it reduces the efficiency of visual recognition; in contrast with normal perceivers, some patients with integrative agnosia identified silhouettes better than line drawings, whose internal details apparently led to incorrect segmentation. Figure-ground segregation and other forms of perceptual grouping, such as deriving a unified whole based on Gestalt heuristics, may be especially challenging for such individuals. The rapid and efficient access to the local information may impede the patients' ability to gain access subsequently to the global information.

Visual object agnosia, especially of the integrative or associative type, can be fairly specific to a semantic category, affecting predominantly living or animate objects or affecting mainly recognition of nonliving or inanimate objects with the latter form being the less frequent manifestation. Interesting theoretical exchanges about the type of information (whether functional versus perceptual) tapped for the recognition of living versus nonliving objects as well as questions about differential lesion localization in these two cases abound in the literature.

Prosopagnosia

Individuals with prosopagnosia are unable to recognize individual faces although they are clearly capable of detecting whether a face is present in the input. The disorder may be so profound that the affected individual may fail to recognize himself or herself from photographs and may fail to recognize family members and close friends too. In many instances, these patients can describe the face in detail, including the age and gender of the person, but still be unable to say whose face it is. The ability to extract information about emotional expression is preserved in some cases but impaired in others. Invariably, these patients rely on nonface characteristics such as gait or voice to recognize the individual. The deficit is typically perceptual in nature, rather than arising from a memory (although there are prosopamnesic cases for whom this is the core of the deficit) or semantic deficit, or from a failure to label the face (an anomia). In most cases, the affected individual also performs poorly at discriminating between two faces, even novel faces and even under conditions in which they are given unlimited time in which to make the discrimination decision. Many studies suggest that these patients process faces differently than do controls following the lesion, focusing more on individual features than on the holistic or second-order relations between the features. The lesion that gives rise to prosopagnosia is usually secondary to occipitotemporal lesions, affecting the fusiform and lingual gyri, and the growing consensus is that a unilateral right lesion to these regions is sufficient to give rise to this disorder.

Agnosia for Words

This is also known as pure alexia, alexia without agraphia, or pure word blindness. Although this phenomenon is usually discussed in the context of language impairments, it is an agnosic symptom because subjects who suffer from this deficit show a language impairment limited to visually presented stimuli (e.g., reading), but not to auditorily presented stimuli. These patients may be able to spell words out loud and to write well but fail to read their own handwriting subsequently. This disorder has also been called ventral simultanagnosia to indicate the failure to process multiple letters simultaneously, and this failure manifests most dramatically in word reading where individuals laboriously and sequentially process one letter at a time (giving rise to the label "letter-by-letter" reading). Pure alexia is one of the more common forms of agnosia and can be severely debilitating, with some patients requiring 1.4 seconds to process each letter in a string. Whether this form of agnosia is entirely limited to word recognition or whether other classes of visual objects are also affected, perhaps to a lesser degree, remains controversial. The lesion site is typically in the left ventral occipitotemporal cortex, sometimes, but not always, involving the splenium of the corpus callosum.

Agnosia for Landmarks and Environment

An impairment in recognizing landmarks or scenes, referred to as topographic agnosia, can co-occur with prosopagnosia or it can also occur in isolation. As with other forms of agnosia, these patients perform well on a wide range of memory and basic perception tasks. In these topographic cases, familiar landmarks and buildings lose their familiarity to the patients. Unsurprisingly, these patients get lost and are unable to learn new routes. They show undue reliance on street names and numbers of buildings to identify the landmarks. Topographic agnosia can be differentiated from other disorders affecting spatial orientation in a large-scale environment and a taxonomy. The lesion giving rise to topographic agnosia most commonly follows bilateral or right posterior artery infarction, and the lesions (as in prosopagnosia) may implicate the fusiform and lingual gyri, extending to the parahippocampal gyrus. Neuroimaging studies provide converging evidence for topographic-specific representations in these cortical regions, such as in the so-called parahippocampal place area.

Developmental Forms of Agnosia

The taxonomy and illustration of some forms of agnosia described earlier refer to the object recognition failures in individuals who were premorbidly normal. There is growing interest in some forms of agnosia that appear to be present from childhood. These childhood agnosias may be subdivided into two subgroups: *developmental agnosia* in which individuals are born normal and then become agnosic following a lesion, such as stroke, sustained during childhood, and *congenital agnosia* in which the disorder is present from birth or, at a minimum, lifelong, in the absence of any obvious neurological concomitants. Developmental dyslexia, or at least one subtype thereof, may be a form of childhood agnosia, and although some individuals appear to have an obvious underlying neural concomitant, this is not the case for all individuals. Recently, there have been several reports of individuals who are congenitally prosopagnosic and reports of individuals who are congenitally agnosic for color. Much research remains to be done with the individuals with the congenital variants to characterize the disorder fully and to uncover the underlying neural mechanism that may give rise to these disorders. There also appears to be a familial hereditary aspect associated with these disorders, and this promises a fruitful avenue for further genotype studies in these individuals and their family members.

Progressive Posterior Cortical Atrophy

Just as there appear to be forms of agnosia that arise in childhood, so there appear to be forms of agnosia that can occur relatively selectively as a form of progressive cortical deterioration associated with aging. Posterior cortical atrophy refers to the progressive decline in complex visual processing ability along with relative sparing of other cognitive and perceptual function. It is associated with occipitoparietal atrophy and hypometabolism on single photon emission computed tomography (SPECT) or positron emission tomography (PET) scans and can be independent of Alzheimer's disease. Progressive prosopagnosia is a degenerative disorder in which there is a progressive impairment in the recognition of faces. This syndrome is part of the fronto-temporal dementias (FTDs), which may present as focal atrophy in any combination of the right and left frontal or temporal cortices and is rather uncommon. Other forms of progressive agnosia may include agnosia for words and agnosia for objects but further description of such cases and systematic data collection are required in such cases.

Marlene Behrmann

See also Agnosia: Auditory; Agnosia: Tactile; Face Perception; Neuropsychology of Perception; Object Perception

Further Readings

Aguirre, G. K., & D'Esposito, M. (1999). Topographical disorientation: A synthesis and taxonomy. *Brain, 122*(9), 1613–1628.

Behrmann, M., & Avidan, G. (2005). Congenital prosopagnosia: Face-blind from birth. *Trends in Cognitive Science, 9*(4), 180–187.

Behrmann, M., & Williams, P. (2007). Impairments in part-whole representations of objects in two cases of integrative visual agnosia. *Cognitive Neuropsychology, 24*(7), 701–730.

Farah, M. J. (1991). Patterns of co-occurrence among the associative agnosias: Implications for visual object recognition. *Cognitive Neuropsychology, 8,* 1–19.

Farah, M. J. (2004). *Visual agnosia.* Cambridge: MIT Press.

Teuber, H.-L. (1968). Alteration of perception and memory in man. In L. Weiskrantz (Ed.), *Analysis of behavioral change.* New York: Harper & Row.

AIR QUALITY

Residents who live near refineries or pig farms may be exposed to air that is of reduced quality, a situation that will rarely go unnoticed. They may perceive the air as unpleasant, unacceptable, and not appropriate for inhaling—all of which connotations are implied by the term *quality*—and they are well equipped by their senses to make such judgments. This entry provides an outline of the sensory systems and measurements involved in air quality evaluation, the health effects of exposure to bad air, and guidelines established to protect humans from exposure to bad air.

Senses Involved in Air Quality Judgment

Our sense of olfaction contributes to our evaluation of air quality by imparting sensations of intensity, (un)pleasantness, character (e.g., whether the odor smells solvent-like or putrid), and identity ("This is diesel fuel"). In addition, many airborne chemicals give rise to feelings of burning and stinging in the eyes, nose, or upper airways associated with the *common chemical sense (chemesthesis).* Both senses contribute to a person's evaluation of air quality. For example, air quality will be judged as bad when a person smells ammonia while feeling a sting up the nose during work.

Health Effects and Contributing Factors

We distinguish between indoor air quality and outdoor air quality. The air quality out of doors can be bad near highways, in industrial areas, and near animal operations. Indoor air quality may be jeopardized in offices as a result of bad ventilation, in the workplace when working with organic solvents, or at home and in schools because of fumes from carpet glue or building materials.

Depending on the chemical and its concentration, people may experience health effects from exposure such as sensory irritation of the eyes, nose, and throat; respiratory problems; or even central nervous system effects such as losing consciousness. Exposure to toxic substances at work can damage the olfactory epithelium—where the odor receptors are located—and diminish a person's sense of smell. On the other hand, decreased odor sensitivity to chemicals as a result of occupational exposure may also be the result of adaptation. Often, the situation reverses after exposure has discontinued.

We speak of *direct* health effects, such as described earlier, when people report health effects that are in line with toxicological models concerning the relation between health effects and chemical concentration. Yet, complaints about air quality and associated symptoms cannot always be explained scientifically. This may be the case when chemical concentrations cannot be measured with detection instruments, or are too low to account for the reported effects. This may be explained by individual variability in sensitivity because some people can pick up certain smells at much lower concentrations than other people can—or detection instruments—and yet others are sensitive because of allergic conditions or asthma.

Health effects may also be *indirect* consequences of chemical exposure, in the sense that the effects are mediated by the perception or interpretation of the odor. For example, when people smell an unpleasant odor near a factory, they may interpret that odor as a warning signaling harm, and react accordingly. Such reactions may involve a stress response, which causes physiological changes in the body leading to increases in blood pressure, heart rate, and sweating. In this manner, health effects that were really brought on by stress associated with perception of odor as threat can be mistakenly attributed to the chemical composition of the odorous air.

To better understand how air quality is perceived in the latter scenario, psychological factors need to be considered. Many studies have focused on annoyance experienced from exposure to malodorous air as an indicator of bad air quality. Being worried about the environment, or potential health effects, as well as endorsing certain beliefs such as that unpleasant odors can threaten your health, contribute to this annoyance. Furthermore, intensity of the odor, frequency and duration of exposure, and unpleasantness are all predictors of annoyance.

The discrepancy between subjective perceptions of air quality and objective air quality can become quite extreme, as in *multiple chemical sensitivity (MCS)*. MCS is a syndrome of still unexplained causes, characterized by subjective symptoms such as nausea, muscular pain, memory difficulties, and fatigue following low levels of chemical exposure.

Measuring Air Quality and Setting Standards

To protect workers from adverse health effects related to bad air quality at work, occupational exposure limits (OELs) have been formulated by government agencies and professional organizations. These define the highest levels of chemicals to which a human may be exposed without experiencing adverse health effects. For a number of substances, irritation of the nose, throat, and eyes provides the base for the OEL, whereas for some other (obnoxious) odors, odor effects provide the base. To assess the levels where chemical substances start to irritate or smell, human odor or irritation detection thresholds (i.e., how strong the chemical needs to be before it is noticeable as either odor or irritation) have been determined using *olfactometry*. Irritation thresholds can also be measured directly at the eyes instead of the nose. More objective assessments of sensory irritation include the electrophysiological *nasal mucosal potential,* which refers to the summed electrical response of the trigeminal chemoreceptors in the nasal mucosa, or ocular hyperemia (“red eye”), which can be assessed by comparing photographs of the eyes pre- and post-exposure.

Standards for outdoor air typically aim to keep annoyance from odor at acceptable levels. Annoyance and related complaints are assessed via surveys or resident diaries in the area as indicators of effects of odor exposure. Olfactometric assessments of threshold or pleasantness evaluations of air sampled near the source and evaluated by a trained panel of judges in a lab, or by a “sniffing panel” at various distances from the source itself, are examples of methods used to index odor dose. Formulating guidelines employing dose-effect relationships based on the previous, while considering economical and financial interests of the involved parties, can present a major challenge to the (local) governmental bodies involved in standard setting.

So far, emphasis has been on perception of bad air quality to alert a person to a health hazard. However, air quality can also be good and promote health. Lemon scent, for example, has been shown to enhance both positive mood and physiology.

Monique A. M. Smeets

See also Common Chemical Sense (Chemesthesis); Olfaction; Olfactometry; Olfactory Adaptation

Further Readings

Das-Munshi, J., Rubin, G. J., & Wessely, S. (2006). Multiple chemical sensitivities: A systematic review of provocation studies. *Journal of Allergy and Clinical Immunology, 118,* 1257–1264.

Schiffman, S. S., Walker, J. M., Dalton, P., Lorig, T. S., Raymer, J. H., Shusterman, D., & Williams, C. M. (2000). Potential health effects of odor from animal operations, wastewater treatment, and recycling of byproducts. *Journal of Agromedicine, 7,* 1–81.

Smeets, M. A. M., Kroeze, J. H. A., & Dalton, P. H. (2006). Setting occupational exposure limits in humans: Contributions from the field of experimental psychology. *International Archives of Occupational and Environmental Health, 79,* 299–307.

Sucker, K., Both, R., & Winneke, G. (2001). Adverse effects of environmental odours: Reviewing studies on annoyance response and symptom reporting. *Water Science and Technology, 44,* 43–51.

AMBLYOPIA

The most common cause of vision loss in children of the developed world is a developmental disorder called *amblyopia,* sometimes called lazy eye. The term *amblyopia* is used to refer to a heterogeneous group of conditions that present to the ophthalmologist as reduced visual acuity in one eye despite normal ocular health and optimal refractive correction (not correctable with glasses). The fellow (unaffected) eye demonstrates normal clinical visual acuity. Amblyopia usually begins in infancy and becomes acute in childhood. It is associated with multiple conditions, such as a monocular cataract, or more commonly a deviated eye or a difference in near- or far-sightedness between the two eyes. The typical result is poor vision in one eye and severely impaired binocular vision. These abnormalities usually persist into adulthood. Many people know someone affected, given its prevalence of 2 to 4% in the general population. This entry describes types of amblyopia, visual loss, etiology, neural substrates, and treatment of amblyopia.

Types

Most specialists distinguish *anisometropic* amblyopia, which refers to conditions of unequal refractive error (focus) between the two eyes, from *strabismus,* which refers to a deviated eye that may or may not involve loss of acuity. Depending whether the eye deviates toward the temple or toward the nose, the strabismic condition is called exotropia or esotropia. In the United States, esotropia is the more common presentation. In the case of exotropia, some subjects adopt the strategy of moving their head to alternate left and right eye viewing, thus acuity is preserved in both eyes, but binocular integration does not develop normally.

Visual Loss

The visual deficits seen in anisometropic and strabismic amblyopia are different in many respects, yet they share some common characteristics. Numerous studies have compared the monocular performance of amblyopic and fellow eyes in both subtypes. The impaired eye in both types of amblyopia generally exhibits decreased visual acuity and reduced contrast sensitivity (particularly for fine patterns). However, strabismic amblyopes can be distinguished from the anisometropic subtype by an additional loss of accuracy for position judgment tasks—for example, judging the alignment of two lines (called Vernier acuity). Other tasks with stimuli containing multiple targets demonstrate crowding effects in strabismic observers. In other words, a target that would be visible when presented alone can no longer be discerned in a crowded array of targets because of interference from nearby elements. Moreover, strabismic observers sometimes report geometrical distortions in their perceptions. All these additional deficits support the idea that strabismus may cause a different type of amblyopia than does anisometropia, with greater spatial localization impairments. Finally, both types of amblyopia show deficits in tests that require binocular integration such as depth perception, with the strabismic subjects generally performing worse than anisometropic. In particular, the type of depth perception that integrates the small differences in the left and right eye's image (called stereopsis) is severely affected.

Etiology

Debate continues about how the strabismic and anisometropic subtypes differ from each other in disease etiology and visual performance. One approach to this abiding question is to consider the relationship between performance on monocular versus binocular tests. Almost all subjects with amblyopia show impairments on binocular tests, but subjects can be categorized based on the degree of residual function. Most subjects with strabismus, and some more severe subjects with anisometropia, are non-binocular by this definition. It has been recently proposed that the degree of impairment in binocular functions may predict the pattern of monocular deficits, and thereby help explain the mechanisms of amblyopia. Specifically, suppression of the amblyopic eye's inputs by the fellow eye's inputs may be an important etiological factor in the progressive development of amblyopia in non-binocular subjects. This may be an adaptive response to avoid double vision, but it leads to loss of binocular integration, and monocular impairments for spatial localization abilities such as Vernier acuity.

Neural Substrates

General consensus is that the cerebral cortex is the major site of neural changes in amblyopia, historically based on animal models. These models generally induce an amblyopia-like condition through monocular enucleation, lid suture, blurring, or eye deviation during development. In general, impaired vision in one eye leads to abnormalities in the visual system of the brain because the weak eye fails to form normal neural connections that allow it to cooperate with the dominant eye. Specifically, these models show reduced visual acuity and contrast sensitivity in the physiological response of individual neurons in the primary visual cortex (V1) driven by the amblyopic eye. They also show decreased numbers of binocular or disparity-sensitive neurons. Moreover, there is some physiological evidence for active suppression of the strabismic eye by the fellow eye in V1 binocular neurons. Nevertheless, the overall extent of physiological deficits measured to date in V1 is not sufficient to account quantitatively for the visual performance losses observed.

Growing evidence indicates that dysfunction at higher-levels of the visual cortex is important for the neurological account of amblyopia. Currently, this evidence comes from animal studies, from human visual performance, and especially from human brain imaging techniques. Limited information is available about higher-level cortical visual areas in animal models, but deficits have been generally reported. In studies of visual performance, higher-order impairments have been discovered that relate to global form and motion perception, and are not likely to be a simple consequence of neural deficits at the earliest stages of the visual system. During the past 10 years, anatomical brain imaging of children with acute amblyopia as well as functional brain imaging of metabolism levels of adults with long-standing amblyopia has been possible. The results clearly indicate that both the primary visual cortex and higher-level cortex in the parietal and temporal regions are abnormal in subjects with amblyopia.

Treatment

Current strategies for treating amblyopia rely on occluding (patching) the fellow eye, encouraging a type of passive visual training with the amblyopic eye. These therapies can be moderately successful in improving acuity in the weak eye under optimal conditions, when the amblyopic deficit is small and treatment is started early in life. However, patient compliance is often poor, and binocular integration often remains impaired. New approaches are needed to pioneer more effective therapies. Finally, the advent of modern human genetic analysis is relevant to amblyopia because a genetic component to amblyopia is supported by twin studies and other analysis. The potential identification of specific genes would benefit our understanding of the etiology and subtypes of amblyopia in the future.

Janine D. Mendola

See also Brain Imaging; Depth Perception in Pictures/Film; Experience-Dependent Plasticity; Eye: Structure and Optics; Low Vision; Neuropsychology of Perception; Visual Acuity

Further Readings

Conner, I. P, Odom, J. V., Schwartz, T., Mendola, J. D. (2007). Retinotopic maps and foveal suppression in visual cortex of amblyopic adults. *Journal of Physiology, 583*(1), 159–173.

Kiorpes, L., & McKee, S. P. (1999). Neural mechanisms underlying amblyopia. *Current Opinion in Neurobiology, 9*(4), 480–486.

McKee, S. P., Levi, D. M., & Movshon, J. A. (2003). The pattern of visual deficits in amblyopia. *Journal of Vision, 3*(5), 380–405.

AMERICAN SIGN LANGUAGE

American Sign Language (ASL), described in this entry, is the visual-motor language used among members of the deaf community in the United States. ASL is one of many different sign languages used by deaf persons around the world. In most countries, deaf persons have their own distinct sign language. Once ASL and other signed languages were recognized as true languages, investigators began examining how the modality difference between spoken and signed languages affected language acquisition, language structure, and brain processing.

For centuries, the sign languages used by deaf persons were not accepted as true languages and often were depicted as consisting primarily of pantomimic gestures. Such pantomimic gestures or "pictures in the air" were not seen as constituting a valid symbol system. Although some ASL signs are pantomimic and clearly resemble the objects, actions, or properties for which they stand, the meanings of most ASL signs are difficult for nonsigners to guess correctly. Rather, the meanings associated with most ASL signs depend on the usage of these signs within the signing American deaf population.

Today, most scholars accord full linguistic status to ASL and other sign languages. The person most responsible for changing people's views about the linguistic basis of sign languages was William Stokoe. Stokoe's pioneering research showed that ASL signs had a distinct linguistic structure and that this structure was different from that of English and other spoken languages. After examining the formational structure of numerous ASL signs, Stokoe proposed that three aspects of sign formation distinguished any one sign from another. These three formational elements were the place or location where a sign was made, the shape or configuration of the hand(s), and the action or movement of the hand(s) and arm(s) while making the sign. Altogether, Stokoe identified 55 different locations, handshapes, and movements from which all ASL signs were composed. These different formational elements operate in a manner largely analogous to phonemes in spoken languages.

Studies of the historical origins of ASL have shown a strong influence of French Sign Language. This influence is related to the establishment of public education programs for deaf students. In 1815, American Thomas H. Gallaudet traveled to Europe to learn about educational programs for deaf students. He returned to the United States with Laurent Clerc, a highly talented teacher from the school for deaf students located in Paris. In 1817, Gallaudet and Clerc helped establish the first public school for deaf students in the United States, now known as the American School for the Deaf. As a prominent teacher there, Clerc relied heavily on his knowledge of French Sign Language. Many of the signs used at the American School were widely disseminated when pupils went home and teachers moved to other newly established schools for deaf students. Some ASL signs, moreover, appear to stem from the indigenous sign-communication systems present among deaf persons in the early United States. Finally, because ASL is a vibrant, living language, new signs frequently are added to its lexicon as the need arises.

ASL and other signed languages may be learned at different ages and in different ways than spoken languages. The hearing status of the parents often affects how children acquire ASL. Most children (hearing or deaf) of deaf parents learn to sign much like hearing children of hearing parents learn to speak. These deaf parents model ASL usage and engage their children in conversation using ASL. Most deaf children, however, are born to hearing parents. Because these parents traditionally do not sign, their children often do not learn ASL until they attend a school for deaf students or join a deaf community organization. This process appears to be changing as more hearing parents and teachers learn to sign.

Because ASL and other sign languages are visual-motor languages, they often transmit information in quite different ways than auditory-vocal languages. ASL makes extensive use of visual and spatial relationships in its transmission of information. Changes in the direction of sign movement, the speed of sign production, the size of signs, and the location of signs in space affect the meaning of sign utterances. Some ASL verbs, for example, vary the direction of their movement to indicate who is performing the action, who the recipient of the action is, and where the action takes place. English, in contrast, relies quite heavily on word order and prepositions to convey such relationships. Given these differences in how information may be effectively transmitted in ASL and in English, it should not be surprising to learn that their grammatical systems are quite different.

The difference in modality between spoken and signed languages has led investigators to ask whether sign languages are processed in the same areas of the brain as spoken languages. Studies of persons who have experienced an aphasia or loss in language abilities (typically after suffering a stroke) provide one source of information on this subject. Damage to particular areas of the left hemisphere is related to certain forms of spoken language impairments; deaf persons with sign language impairments also show evidence of left-hemisphere damage in these same (or closely related) areas.

Although studies of aphasias in deaf and hearing persons have underlined important similarities in left-hemisphere processing for signed and spoken languages, more recent neural imaging studies of fluent ASL users have shown substantial right-hemisphere involvement as well. This right-hemisphere activation in ASL users may reflect the visuospatial processing present in signed languages but not involved in oral languages. Therefore, the answer to the question about whether signed and spoken languages are processed in the same areas of the brain is a complex one: they both show strong similarities in left-hemisphere processing, but the unique characteristics of signed languages such as ASL may activate right-hemisphere areas that are not required for spoken language processing.

Today ASL is seen as a full language capable of effectively conveying a wide range of information. How ASL transmits this information and where ASL is processed in the brain have been rich areas of research in recent years.

John D. Bonvillian

See also Aphasias; Audition: Disorders; Brain Imaging; Cortical Reorganization Following Damage; Language

Further Readings

Bonvillian, J. D. (1999). Sign language development. In M. Barrett (Ed.), *The development of language* (pp. 277–310). Hove, East Sussex, UK: Psychology Press.

Newman, A. J., Bavelier, D., Corina, D., Jezzard, P., & Neville, H. J. (2002). A critical period for right hemisphere recruitment in American Sign Language processing. *Nature Neuroscience, 5,* 76–80.

Stokoe, W. C., Jr., Casterline, D. C., & Croneberg, C. G. (1965). *A dictionary of American Sign Language on linguistic principles.* Washington, DC: Gallaudet College Press.

Wilbur, R. B. (1987). *American Sign Language: Linguistic and applied dimensions* (2nd ed.). Boston: College-Hill Press.

AMES DEMONSTRATIONS IN PERCEPTION

The Ames demonstrations consist of about 25 laboratory set-ups that were designed and constructed, from about 1934 to 1950, by a U.S. artist and optical physiologist named Adelbert (Del) Ames, Jr. They became commonly known as the *Ames demonstrations in perception* in the early 1950s, when one of Ames's associates, perceptual psychologist William H. Ittelson, published a users' manual (with instructions on how to rebuild them) with that phrase as its title.

Although there has never been consensus about the validity or significance of these demonstrations, they are most often said to show the ambiguity of retinal images when an object is perceived from a single static point of view. It is also often claimed that they show that the process of seeing is not merely the passive reception of facts, but an interactive (or constructive) process by which we arrive at conjectures, based on retinal data, past experience, expectations, and other factors. This entry describes the Ames demonstrations in perception.

Background

To understand the Ames demonstrations, it is helpful to know the circumstances that contributed to their development. Early in his career, Ames studied law at Harvard University, where one of his influences was philosopher and psychologist William

James, a leading proponent of pragmatism. After practicing law only briefly, Ames turned instead to the study of art, and for several years, collaborated with his sister, Blanche Ames Ames, a painter and illustrator whose husband was a scientist. In the process, Ames's interests shifted toward scientific methods of painting more "realistic" scenes, a quest that eventually led him to the study of optical physiology, and, in time, to join the research faculty at Dartmouth College. There, in 1928, he discovered a visual dysfunction he called *aniseikonia,* a condition in which a person's two retinal images differ enough in size and shape that the brain cannot easily fuse them as one. Some of the Ames demonstrations (which he invented later) may have been prompted by this discovery, partly because those with aniseikonia are susceptible to spatial distortions: Flat surfaces tend to look tilted, rectangles look trapezoidal, and trapezoidal shapes look rectangular.

The Demonstrations

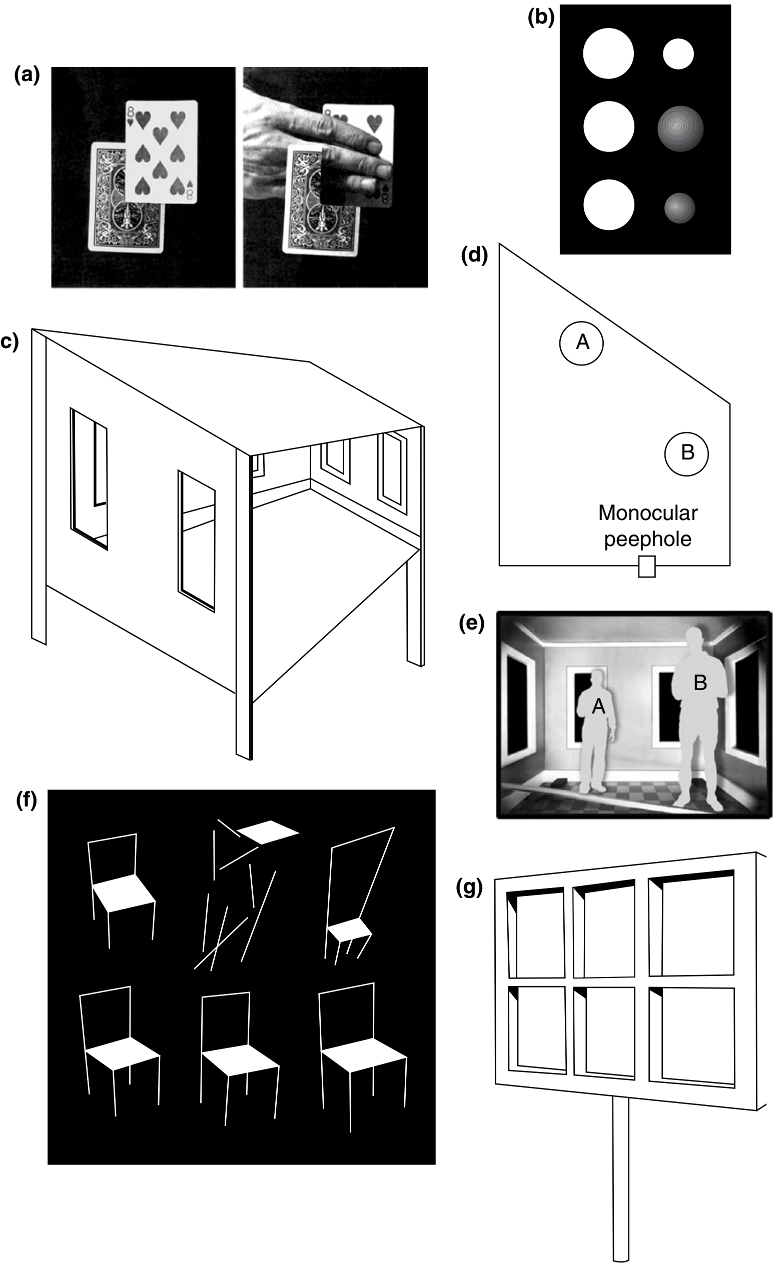
As early as 1934, Ames designed his first "distorted room" in which, from a designated viewing point, trapezoidal walls appear to be slanted rectangles (Figure 1, c–d). His initial room was monocular, requiring that it be viewed through a one-eyed peephole. From that static point of view (and from no others), the space inside appeared to be a conventional room interior, with a checkerboard floor and three visible walls, with a pair of rectangular windows on each.

However, when other components were introduced into the space, the results were astonishing. If two people of the same height stood in opposite corners, one appeared to be tall, the other short; and if they exchanged positions, one appeared to grow, the other to shrink. A ball placed in a wooden trough appeared to roll uphill. Liquid poured from one container to another appeared to be strangely inclined to the side. And if an observer was given a lecturer's pointer and instructed to touch the room's back corners (while looking through the peephole), the result was a baffling sensation in which one's optical reality was at odds with one's kinesthetic reality.

The scientific reason for these "illusions" was the calculated distortion of the room, which is obvious when it is viewed from any point other than the peephole. Supportive of that, the right wall (in all its aspects: the windows, the molding, the checkerboard floor) is only one half the dimension of the left wall, while the rear wall is a trapezoid, inclined at an obvious angle. The floor and ceiling are also inclined and distorted. Nevertheless, from the single position of the peephole, the shape of the room appears normal, so that, with that as the standard, things brought into the room will most likely look oddly misshapen.

People are generally taken aback when they first experience the Ames distorted room, partly because, even when they know in advance that the room is distorted, its effects may be no less convincing. The same is also generally true of Ames' other demonstrations, which he gradually developed in the remaining years of his life. Among the most interesting are the following:

- The *overlay* demonstration (Figure 1a) in which portions of certain playing cards were cut out and arranged to look as if they were in front of (and overlapping) other cards, when in fact they were behind.
- The *size-brightness* demonstration (Figure 1b) in which balloons were made to look nearer or farther away not by changing their actual distance from the viewer, but by precisely modifying their size and illumination.
- The *chair* demonstration (Figure 1f), invented in 1945, was installed in a large wooden box with three peepholes, spaced equally on the front. Looking through each peephole, one sees what seems to be a chair of the same size, in the same position (as shown on the bottom row of the illustration). However, in opening the lid and looking inside the box (as shown on the top row), it is evident that only one of the three is the customary shape of a chair, whereas the other two are constructions of wires, rods, and other shapes.
- The *rotating trapezoid window* (Figure 1g), which Ames invented in 1947, is a flat piece of sheet metal, in the shape of a trapezoid and deceptively painted with shadows, that appears to be a normal dimensional window in perspective. Mounted on a motorized vertical shaft, as the window rotates, it appears instead to oscillate (to sway back and forth). When other components are suspended from it, they behave in unbelievable ways.



Figures 1 Views of Five of the Laboratory Set-Ups (Developed by Adelbert Ames II) That Are Commonly Referred to Now as the Ames Demonstrations in Perception

Source: Illustration © by Roy R. Behrens, 2008. Reprinted with permission.

Notes: (a) The *overlay* demonstration, in which a playing card, from which a portion has been cut out, is positioned in front of a second card, making the furthermost card appear instead to be in front. (b) The *size-brightness* demonstration in which two balloons appear to move closer or further away by changes in size and illumination. (c–e) The *distorted room* in which people and commonplace objects are viewed inside a room that seems (from one point only, through a monocular peephole) to be a simple, normal room. Seen from any other angle, it is apparent that the right wall is one half the size of the left wall, the rear wall is a trapezoid, and both the ceiling and the floor are oddly shaped and slanted. (d) A bird's-eye view indicates the position of the peephole, and (e) the physical locations of two people of comparable height. (f) In the *chair* demonstration, three forms viewed through peepholes appear at first to be chairs (bottom row), but look dramatically different from a higher alternative angle (top row). (g) In the *rotating trapezoid* demonstration, a trapezoidal shape that has been painted to resemble a perspective view of a window, and mounted on a motorized shaft, appears not to rotate but to sway back and forth.

Anamorphosis

The Ames demonstrations were not unprecedented, in the sense that they have much in common with an historic artistic distortion technique called *anamorphosis*, and to other perspective illusions employed in the design of theatrical sets.

In anamorphic constructions, the image looks distorted when viewed frontally (as is customary) but correctly proportioned when seen from the side (often indicated by a peephole). For example, Ames was well acquainted with the work of scientist and philosopher Hermann von Helmholtz, who more than 50 years before had noted that an infinite variety of distorted rooms could be devised that, from a monocular peephole, would nonetheless seem to be normal.

Far in advance of Helmholtz, this same kind of visual distortion was used as early as 1485 by Leonardo da Vinci (and probably even earlier by Chinese artists) as an offshoot of perspective. As early as the 16th century, a group of Dutch artists made anamorphic "peepshows" called *perspectyfkas* or perspective cabinets, a few of which still exist and are now on display in museums. These artistic constructions have much in common with the Ames distorted room, the chair demonstration, and the rotating trapezoid window. In fact, many of the Ames demonstrations make use of trapezoids that appear rectangular, tilted surfaces that appear to be flat, or flat surfaces that appear tilted.

Most of the Ames demonstrations require (or at least work optimally with) a monocular peephole. This is because they rely on the fact that any number of external constructions could produce the same retinal image. Ames explicitly stated that (as Helmholtz had noted) he could have constructed an infinite number of distorted rooms, no two of which would have the same physical shape, yet each would appear to be normal. In the end, he only constructed a few.

Implications

The Ames demonstrations might never have become well known had they not been embraced and promoted by other people who saw them as palpable evidence of their own convictions. They were of particular value to proponents of transactional psychology (not to be confused with transactional analysis), a 1950s spin-off of pragmatism that was partly inspired by the "transactional approach" of philosopher John Dewey (who first saw the demonstrations in 1946, and then corresponded with Ames until 1951). Ames believed, as Dewey did, that we are not passive recipients of a given reality, but instead are active participants in a give-and-take exchange (a "transaction") in which split-second assumptions are made about the nature of reality.

The demonstrations have also had lasting effects on other aspects of culture. Even today, one or more of the demonstrations are invariably mentioned in textbooks on perception, and it is not uncommon for one or more to appear in television documentaries, video clips, cinematic special effects, or advertising commercials.

Roy R. Behrens

See also Magic and Perception; Object Perception; Pictorial Depiction and Perception

Further Readings

Behrens, R. R. (1994). Adelbert Ames and the cockeyed room. *Print, 48,* 92–97.

Behrens, R. R. (1998). The artistic and scientific collaboration of Blanche Ames Ames and Adelbert Ames II. *Leonardo, 31,* 47–54.

Behrens, R. R. (1999). Adelbert Ames, Fritz Heider, and the chair demonstration. *Gestalt Theory, 21,* 184–190.

Cantril, H. (Ed.). (1960). *The morning notes of Adelbert Ames, Jr.* New Brunswick, NJ: Rutgers University Press.

Ittelson, W. H. (1968). *The Ames demonstrations in perception.* New York: Hafner.

Kilpatrick, F. P. (Ed.). (1961). *Explorations in transactional psychology.* New York: New York University Press.

AMODAL PERCEPTION

Amodal (meaning "without" modality) *perception* is perception of information that is common or redundant across multiple senses (e.g., auditory, visual, tactile). Amodal information includes changes along three basic parameters of stimulation—time, space, and intensity. Properties of objects and events such as temporal synchrony, rhythm, tempo, duration, intensity, and co-location are common across auditory, visual, and proprioceptive stimulation. Properties such as shape, substance, and texture are common across visual and tactile stimulation. For example, the same rhythm and tempo can be detected by seeing or hearing the pianist strike the notes of the keyboard, and the same size, shape, and texture can be detected by seeing or feeling an apple.

Virtually all events occur across time, are distributed across space, and have a characteristic intensity pattern, so virtually all events provide amodal information. For example, speech comprises changes in audiovisual synchrony, tempo, rhythm, and intonation (intensity changes) that are common to the movements of the face and the sounds of the voice. Self-motion produces proprioceptive feedback (information from the muscles, joints, and vestibular system) that is synchronized and shares temporal and intensity changes with the sight of self-motion (e.g., seeing and feeling one's hand move). Perceiving amodal information is critically important for organizing early perceptual and cognitive development and for accurate perception of everyday events in children and adults alike.

The term *amodal* has also been used in a different sense—to refer to perception in the absence of direct information from a specific sense modality. For example, in visual perception, *amodal completion* describes how we perceive a unitary shape (e.g., a ball), even when part of the object or shape is occluded (hidden) behind another object (e.g., a block). Even infants can accurately perceive a partially hidden shape if the occluder is moved back and forth, progressively revealing and then hiding the object's contours. Scientists propose that we perceive unitary shape by detecting visual invariants (patterns that remain constant across change) through object motion, whereas others maintain that we must fill in the missing information by inference or cognitive processes. Whatever the process, the term *amodal* referring to incomplete information is not consistent with the previous definition (which refers to information that is fully available and can be directly perceived through more than one sense) and, thus, will not be discussed further. This entry describes the history, theory, and development of amodal perception.

History and Theory

For centuries, philosophers and scientists have been intrigued by how we perceive unified objects and events even though our senses provide specific information through separate sensory channels. How are these different sources of stimulation bound together? Further, why do our senses provide overlapping and redundant information for many qualities of objects? The concept of amodal perception addresses these important questions and dates back more than 2,000 years to the time of Aristotle. Aristotle proposed a *sensus communis* (an amodal or common sense) that detected qualities that were common to several senses. These common sensibles included number, form, rest, movement, magnitude, and unity—information that today is considered amodal.

Centuries later, philosophers such as John Locke and George Berkeley took a different approach to the question of perceiving object and event unity. They proposed that sensations had to be interpreted and integrated across the senses before a person could perceive meaningful objects and events. Until recently, developmental psychologists, including Jean Piaget, thought this process of integration developed gradually through experience with objects. By coordinating and associating what one sees with what one feels and hears, one could construct a coherent, three-dimensional world of objects and events.

This constructivist view was not seriously questioned until James J. Gibson's ecological view of perception emerged in the 1960s, and a view more consistent with that of Aristotle's reemerged. Gibson proposed that the different forms of stimulation from the senses were not a problem for perception, but rather provided an important basis for perceiving unitary objects and events. Our senses, he proposed, work together as a unified perceptual system to pick up information that is invariant or common across the senses—that is, amodal information. If we attend to amodal information, then there is no need to learn to integrate stimulation across the senses to perceive unified objects. Temporal synchrony (the most basic form of amodal information) has been described as the glue that binds stimulation across the senses. For example, by attending to synchrony, the sounds and sights of a single person speaking would be perceived as united. Sights and sounds that are perfectly synchronized belong together and constitute unitary events. Detecting this information prevents the accidental association of unrelated but concurrent sensory stimulation.

Development

Researchers now know that even very young infants are skilled at detecting amodal information, including temporal synchrony, tempo, rhythm, intensity changes, shape, texture, substance, and prosody in speech. Amodal information is highly salient and directs attentional selectivity, for both humans and animals, especially during early development. When videos of two different events are superimposed, playing a synchronous soundtrack to one of them allows even infants to selectively attend to the synchronous event while effectively ignoring the asynchronous one. This attentional salience may be partly the result of the enhanced neural responsiveness generated by redundantly presented amodal information compared with that observed for each modality alone.

The development of a surprising variety of skills depends on the detection of amodal information. Shortly after birth, infants move their eyes in the direction of a sound, coordinating audible and visible space. This provides a basis for detecting further amodal information common to the sight and sound. By 2 to 5 months, infants detect a variety of amodal aspects of objects and events. For example, they detect temporal synchrony, rhythm, and tempo uniting the sights and sounds of objects banging against a surface, as well as more refined temporal information revealing the substance (rigid and elastic) and composition of objects. Infants detect voice-face synchrony in speech by 2 months and later can use it to separate one speech stream from another concurrent one. Even information for emotion is detected by 5 months and is largely amodal, deriving from differences in the timing and intensity of movement and sound. Learning about the self also depends on detecting amodal information for self-motion. By 3 to 5 months, infants detect the congruence between the proprioceptive feedback from their own motion and the visual experience of that motion (e.g., by feeling and seeing their own legs move), and this provides an important basis for separating the self from other individuals. Even maintaining an upright posture requires detecting amodal information common to the visual flow and proprioceptive feedback from body motions. Young infants also detect the common shape, texture, and substance across tactile and visual exploration, allowing them to visually select an object they have previously explored only tactually. Amodal information can also be used to create the illusion of unity as in the *ventriloquism effect.* The ventriloquist creates amodal information by moving the puppet's mouth in time with his own speech sounds and can therefore fool the audience into perceiving that the puppet is speaking. Thus, amodal information simplifies and organizes incoming stimulation, providing a basis for perceiving unitary, multimodal events rather than a "blooming, buzzing confusion" of unrelated sights, sounds and tactile impressions.

Lorraine E. Bahrick

See also Attention: Cross-Modal; Cross-Modal Transfer; Multimodal Interactions: Tactile–Auditory; Multimodal Interactions: Visual–Auditory; Multimodal Interactions: Visual–Haptic; Object Perception; Perceptual Development: Intermodal Perception; Vision

Further Readings

Bahrick, L. E., & Lickliter, R. (2002). Intersensory redundancy guides early perceptual and cognitive development. In R. Kail (Ed.), *Advances in child development and behavior* (Vol. 30, pp. 153–187). New York: Academic Press.

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development.* New York: Oxford University Press.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin, 126,* 281–308.

Lewkowicz, D. J., & Lickliter, R. (1994). *The development of intersensory perception: Comparative perspectives.* Hillsdale, NJ: Lawrence Erlbaum.

ANIMAL CHEMICAL SENSITIVITY

Humans use their chemical senses in many facets of their lives, yet their abilities pale in comparison with those of most animals. Everyday demonstrations of the chemosensory prowess of animals include dogs tracking invisible chemical trails, mosquitoes homing in on human meals, and game fish lured to baited hooks. The chemicals, sensors that detect them, neural pathways that process them, and roles in the life history of animals are as diverse as are the animals themselves. This entry explores the nature and biological basis of such animal chemical sensitivity.

What Is the Chemical World of Animals?

Jakob von Uexküll, in the early 1900s, emphasized that the "umwelt," or sensory world, of each animal species is unique to it, being shaped through evolution to include those stimuli important to that species. Animals live everywhere, from the bottom of the deepest oceans, to the frozen polar caps, to the bowels of other animals, and each is adapted to detect stimuli that are present in that habitat and are of ecological significance to that species. Chemical information dominates the umwelt of most animals. Chemicals called pheromones reveal the nature of members of their own species, including sex, age, social status, health, and individual identity. Chemicals also inform the identity and location of food, shelter, home, predators, competitors, and more. The molecules that convey chemical information are diverse—they can be protons, such as sour-tasting acids in unripe fruit, or they can be macromolecules, such as sweet-tasting proteins in some ripened fruit.

Many animals are much more sensitive to odorants than are humans. Dogs and cats can detect some odorants at thousands of times lower concentrations than humans can. Although the olfactory systems of dogs, cats, and humans have a similar basic organization, the greater sensitivity of dogs and cats comes from their having many more olfactory receptor neurons in their noses. Moths have different "noses" than humans or dogs do, but their olfactory sensitivity to some odorants such as pheromones is also extremely high compared with that of humans.

However, some animals lack sensitivity to specific chemicals that are a central part of the sensory world of humans. Cats lack a sweet taste, a result of a loss of functional receptors in their mouths that detect sugars. Given cats' carnivorous diet, a lack of sweet taste is not a significant deficit.

High sensitivity of animals to particular chemicals is a proven or promised method of controlling populations of species that are pests to humans. Use of specific blends of pheromones in traps is a standard method of controlling some insect species. A similar application holds promise for an aquatic pest, the lamprey.

How Are Chemicals Detected?

Animal chemosensors are diverse in their structure and organization. This reflects differences in their phylogeny and life history. Sensors include nose and tongue of mammals, antenna of insects and crustaceans, pectin of scorpions, tentacles of sea anemones, osphradium of snails, and olfactory pits of squid. Terrestrial animals have olfaction and taste as their two major chemical senses, and these are distinguishable by the chemicals' carrier medium: olfaction uses volatile molecules delivered in air, and taste uses water-soluble molecules delivered in liquid. Aquatic animals mostly use water-soluble molecules, thus raising the question, can olfaction exist underwater? The answer is yes, as exemplified by frogs, which use their "nose" both underwater and in air. Chemical sensors, with their vast diversity, go beyond the dichotomy of olfaction and taste. Fish such as catfish have an olfactory organ; gustatory organs in the form of taste buds inside their mouth, on their barbels, and on their entire body surface; trigeminal chemosensors; and solitary chemoreceptor cells. Crustaceans have many types of chemosensors based on structural organization and innervation, even on a single antenna, mouthpart, or leg. Even the nose of a rodent has multiple chemosensory pathways that differ in structure and function, including the main olfactory epithelium, vomeronasal organ, trigeminal system, septal organ, and Grueneberg ganglion.

Animals have evolved special means of acquiring chemical stimuli and delivering them to the molecular binding sites to initiate the sensory process. Sniffing in mammals, pumping in the olfactory cavity of fish, flicking of antennules and beating of fan organs in crustaceans, and vibrating of wings and antennae of insects all enhance chemical sampling. Once in the nose, volatile chemicals bind to odorant binding proteins, which solubilize and deliver the chemicals to receptors.

Chemoreceptor cells from different phyla have diverse cellular origins, including neurons and modified epithelial cells. Chemoreceptor cells may be packaged into units, such as taste buds of vertebrates or sensilla of arthropods, the latter being porous extensions of the exoskeleton that allow chemical stimuli to reach the chemoreceptors within them while providing protection. Chemoreceptor cells across the animal phyla transduce chemical information into electrical signals using similar principles and molecular components. For example, many animals use metabotropic signaling pathways that contain receptors coupled to intracellular second messenger pathways for amplification of signals. But the components of these transduction cascades can vary across animals.

A common principle for coding of olfactory information, including in some vertebrates, arthropods, and mollusks, is to have the projections of receptor cells and the brain interneurons receiving those projections organized into units, called glomeruli. The "olfactory brains" are composed of a few dozen to thousands of these glomeruli, the number depending on the species. Each glomerulus is activated by a subset of odorants, some odorants being more effective than others, and each odorant activates more than one glomerulus, some more than others. Thus, each odorant is discriminated based on the distinctive activation pattern generated across the glomeruli because of that odorant's unique molecular structure.

Charles Derby

See also Common Chemical Sense (Chemesthesis); Ecological Approach; Evolutionary Approach: Perceptual Adaptations; Olfaction; Olfaction: Evolution of; Pheromones; Taste

Further Readings

Ache, B. W., & Young, J. M. (2005). Olfaction: Diverse species, conserved principles. *Neuron, 48,* 417–430.

Derby, C. D., & Sorensen, P. W. (2008). Neural processing, perception, and behavioral responses to natural chemical stimuli by fish and crustaceans. *Journal of Chemical Ecology, 34,* 898–914.

Dethier, V. G. (1976). *The hungry fly.* Cambridge, MA: Harvard University Press.

Eisner, T. (2003). *For love of insects.* Cambridge, MA: Belknap Press.

Eisthen, H. L. (2002). Why are olfactory systems of different animals so similar? *Brain, Behavior and Evolution, 59,* 273–293.

Li, X., Li, W., Wang, H., Cao, J., Maehashi, K., Huang, L., et al. (2005). Pseudogenization of a sweet-receptor gene accounts for cats' indifference toward sugar. *PLoS Genetics, 1*(1), 27–35.

Sato, K., Pellegrino, M., Nakagawa, T., Nakagawa, T., Vosshall, L. B., & Touhara, K. (2008). Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature, 452,* 1002–1006.

Sorensen, P. W., Fine, J. M., Dvornikovs, V., Jeffrey, C. S., Shao, F., Wang, J., et al. (2005). Mixture of new sulfated steroids functions as a migratory pheromone in the sea lamprey. *Nature Chemical Biology, 1,* 324–328.

Walker, D. B., Walker, J. C., Cavnar, P. J., Taylor, J. L., Pickel, D. H., Hall, S. B., & Suarez, J. C. (2006). Naturalistic quantification of canine olfactory sensitivity. *Applied Animal Behavioural Science, 97,* 241–254.

Zimmer, R. K., & Butman, C. A. (2000). Chemical signaling processes in the marine environment. *Biological Bulletin, 198,* 168–187.

ANIMAL COLOR VISION

Under daylight conditions, most humans experience a richly colorful world in which objects appear to maintain consistent color appearances—such as green grass or blue sky. That familiar association makes it natural to believe that color is an inherent property of objects and lights. Although at first glance this idea seems reasonable, it is wrong. Color is actually a feature of our experience that is constructed from the overall pattern of illumination reaching the eye at any moment as subsequently analyzed and conditioned by the particular details of the organization of the eye and the visual system. Eyes and visual systems show great variation across the animal kingdom, so it is hardly surprising that other animals may experience color in ways that are strikingly different from those familiar to humans. This entry describes how and why color vision varies among the animals.

Critical Definitions and Underlying Biology

Light reaches the eye directly from illuminants, such as a computer monitor, or as reflected from surfaces of objects, such as an apple sitting on the desk. Ignoring changes that may occur across space or over time, light can be physically characterized as varying along just two dimensions—wavelength and intensity. The presence of color vision is formally defined as the capacity of an individual to successfully discriminate between lights or surfaces based on differences in their wavelengths, irrespective of what their relative intensities may be. Note that the definition asks nothing about the quality of what is seen, only whether there is consistent discrimination. If any animal can make such discriminations, it has color vision.

Once the presence of color vision is established, it becomes possible to further characterize the nature of the capacity, for example, by asking what particular wavelengths can and cannot be successfully discriminated. Each of the large range of such tests can provide useful information about the color vision of any given animal. A particularly important description of color vision comes from using a testing procedure called color matching. In such a test, lights of different wavelengths are added together in an attempt to find out which mixtures of lights appear identical to other lights. Although the analysis of such matches can be a bit complex, what results is a description of the dimensionality of color vision. In such an account, if an animal lacks color vision completely, it is characterized as being *monochromatic* (i.e., it experiences a world that is devoid of color variation), an animal with a single dimension of color vision is called *dichromatic,* two dimensions are *trichromatic,* three dimensions *tetrachromatic,* and so on. As the dimensionality of color vision increases, there are corresponding and dramatic increases in the number of separate colors that can be discriminated. For example, it is estimated that a human having dichromatic color vision (a common form of color blindness) can discriminate among some 10,000 surface colors, whereas a human with trichromatic color vision, normal for our species, has the capacity to discriminate something in excess of one million colors.

There is a compelling link between the dimensionality of color vision and the biology of the eye. In the eye, light is absorbed by photopigments that are located in the photoreceptors—in vertebrates, these receptors are called cones. Different types of photopigments vary according to the wavelengths to which they are most sensitive. The number of types of photopigment that are active in daylight viewing is typically directly related to the dimensionality of color vision—monochromats have only a single type of cone pigment, dichromats have two, trichromats have three, and so on. Because of this linkage, it is possible to infer the nature of an animal's color vision from an analysis of the number of types of photopigment contained in eye or, because the photopigment proteins (opsins) are specified by single genes, from direct studies of these opsin genes. Both of these shortcuts allow scientists to learn something about the nature of color vision in a large number of animals for which behavioral tests of color vision may be impractical for one reason or another. This tactic can even be used to infer what color vision must have been like in some ancestral species.

Distribution of Animal Color Vision

What animals have color vision and what is their color vision like? That seemingly simple question cannot as yet be fully answered. A principal problem is that most animal species remain unexamined. That is not surprising because even the total number of animal species remains unknown, with current estimates ranging from perhaps 10 to 100 million species. One useful clue to understanding the extent of animal color vision comes from examinations of the phylogeny of opsin genes. That approach shows that two classes of daylight photopigments (the minimum necessary for color vision) were present early during evolution; certainly at a time well before vertebrates first appeared more than 500 million years ago. It seems likely that color vision has been a persistent feature of animal life over much of the sweep of their history. Beyond that, there have been concrete studies of color vision in enough animal species to permit a number of generalizations.

Invertebrates

This group constitutes the vast majority (about 95%) of all animals. Among the invertebrates that have been studied, there is positive evidence for presence of color vision in many different species—these include familiar animals such as mites, spiders, crustaceans, and a lengthy list of insects (bees, grasshoppers, wasps, butterflies, moths, and various flies). Because of their commercial importance, many good behavioral and biological studies have been done on honeybees. For example, in a behavioral test of color vision, bees might be trained to fly to a small dish of sugar water that is placed on a blue patch of a checkerboard composed of many different colored papers. Over trials, the location of the blue patch is moved about the checkerboard randomly. If the bee consistently finds that blue patch, it must be doing so by using its color vision. Tests like this showed that bees have three types of photopigment and well-developed trichromatic color vision. Early studies led to the surprising discovery that one of three photopigments in the bee has its maximum sensitivity to the ultraviolet (UV) wavelengths, and thus bees can see a part of the spectrum that is invisible to humans. This capacity proves to be highly useful because many flowering plants provide UV signals that bees can exploit in their quests for pollen. Like the bees, many other invertebrates have been found to have UV photopigments, and thus, they too must inhabit a visual world that is quite foreign to humans.

Vertebrates

Although representing only a tiny minority of all animals, there have been good studies of color vision on representative species from each of the large vertebrate groupings. Outside of the mammals, many reptiles, amphibians, fishes, and birds have UV pigments, some having three and others four classes of cone pigments. Many of these species thus have the potential for trichromatic or tetrachromatic color vision and among familiar animals, both goldfish and pigeons have indeed been shown to have tetrachromatic color vision.

Color vision in mammals presents a different picture. Most mammals have only two types of cone pigments, and thus usually have dichromatic color vision. This condition is characteristic of many familiar mammalian species such as dogs, cats, squirrels, and cattle, as well as the close relatives of all these animals. Among the mammals, primates are unique in having added a third type of cone pigment during their evolution and, as a consequence, many primates, for instance humans and their close ape and monkey relatives, have trichromatic color vision. Interestingly, some primate species, among them humans and most of the monkeys found in Central and South America, have striking individual variations in color vision. In humans, such variations include the familiar cases of red-green color blindness, a condition mostly found among men.

Gerald H. Jacobs

See also Color Perception; Color Perception: Physiological

Further Readings

Briscoe, A., & Chitka, L. (2001). Insect color vision. *Annual Review of Entomology, 46,* 471–510.

Jacobs, G. H. (1981). *Comparative color vision.* New York: Academic Press.

Jacobs, G. H. (1993). The distribution and nature of colour vision among the mammals. *Biological Reviews, 68,* 413–471.

Kelber, A., Vorobyev, M., & Osorio, D. (2003). Animal colour vision—Behavioural tests and physiological concepts. *Biological Reviews, 78,* 81–118.

ANIMAL DEPTH PERCEPTION

Cats jump onto tables, horses jump over fences, birds land on window ledges, and chameleons can pick insects out of the air with their long tongues. All of these behaviors, and any others that involve moving around in an environment, require the ability to judge depth and distance. Because depth perception is such an important survival skill, evolution has provided almost every species with a repertoire of ways to evaluate how far away a visual object might be. This entry describes cues for depth, assessing animal depth perception, and animals' use of depth cues.

Cues for Depth

Information about depth and distance is available from several external depth cues in the environment and from internal cues that rely on physiological mechanisms to calculate the position of an object. For the most part, the external cues are referred to as monocular, or pictorial, cues because they can be seen using just one eye, and because they provide the impression of depth that one sees when looking at photographs or paintings. These cues include linear perspective, changes in the density of surface texture, shading, and relative size.

A second class of cues relies on the physiological processing of information created by stimuli that are at different distances. In effect, neurons in the visual cortex compute distance based on the location and movement of images across the retina. Of these physiologically based cues, the most important is retinal disparity, which is the slight difference in the position of images on the retinas of the two eyes that is a consequence of the fact that the eyes are separated on the face. The ability to make use of retinal disparity is known as stereopsis and is the basis for the strong sense of depth that we get when watching a three-dimensional movie wearing the appropriate kinds of glasses.

Assessing Animal Depth Perception

A person can say if he or she sees something in depth and how far away it is. It is much more difficult to measure whether an animal has depth perception, and if so, how good it is. Researchers have used a variety of techniques to assess animal depth perception, each one tailored to a specific species. All share some common principles, however: An animal is taught to make discrimination between two visual stimuli and is rewarded for responding to the correct one. If an animal can make this discrimination, the task is made progressively more difficult until performance falls to chance. This gives a measure of the fineness of the animal's discrimination abilities. For example, several years ago, Donald Mitchell and his colleagues developed the *jumping stand* technique in which kittens jumped from a raised platform onto a surface below. For measuring depth perception, they used two patterned surfaces, one of which was closer than the other. By reducing the difference in distance, they were able to measure the kittens' depth thresholds. For larger animals, such as the horse, a wide board with two trapdoors can be used, and the horse is required to press its nose against the trapdoor to indicate its choice. In an experiment to measure stereoscopic depth perception in falcons, Robert Fox trained birds to fly from a perch to the visual stimuli that were a short distance away.

The Use of Depth Cues by Animals

Most of the monocular depth cues are qualitative; that is, they indicate that one object is closer than another, but they do not give a reliable indication about how much closer. Stereopsis and motion parallax, in contrast, can provide fine depth judgments. So it is of interest to determine which cues animals can use to get a sense of how well they can process depth information. Even the animals with simple visual systems are able to use some depth cues. Male hoverflies can keep themselves at a constant distance from a receptive female until she lands on a flower, and then move in to begin mating. It is likely that these flies maintain their distance by keeping the size of the image of the females on their eyes constant until she stops moving. A more sophisticated approach is taken by locusts, which move their heads back and forth to generate motion parallax. There is only limited information about whether mammals are able to recognize pictorial depth cues, but one study has shown that young monkeys will reach to the apparently closer of two objects in a picture when distance information was provided by pictorial depth cues, and horses seem to be susceptible to illusions that are created by pictorial depth cues.

All species with forward facing eyes that have been tested have shown evidence for stereopsis. This includes species as diverse as insects, amphibians, birds, and mammals. In mammals, many studies in cats and monkeys have shown that they possess disparity sensitive neurons in the visual cortex that respond when a target is located at a specific distance from the fixation point.

Among mammals and birds, many species have eyes located on the sides of their heads. Although this gives them a panoramic view of the world, it does limit the extent to which they can use both eyes to see the same region of the visual field in front of them. Potentially, this could restrict their ability to use retinal disparity as a depth cue. However, it appears that if these animals have even limited binocular fields, in which both eyes can see the same region of space, then they have stereopsis. Horses, for example have their eyes located on the sides of their heads, but do have about 60 degrees of binocular overlap. Tests have shown that they have normal stereoscopic vision and that they are able to recognize depth binocularly when there is no other information available. Only a few studies have attempted to assess stereoacuity, or the accuracy of binocular depth judgments, and it appears that with the exception of monkeys, stereoacuity in mammals is not nearly as good as it is in humans.

Research that has been done on a wide range of species indicates that depth perception is ubiquitous in the animal kingdom. The extent to which a particular animal relies on one cue or another varies considerably, but all animals are able to use the available cues to judge depth and distance.

Brian Timney

See also Animal Eyes; Binocular Vision and Stereopsis; Motion Parallax and Structure From Motion; Retinal Anatomy; Spatial Layout Perception, Neural; Spatial Layout Perception, Psychophysical

Further Readings

Fox, R., Lehmkuhle, S. W., & Bush, R. C. (1977). Stereopsis in the falcon. *Science, 197,* 79–81.

Gunderson, V. M., Yonas, A., Sargent, P. L., & Grant-Webster, K. S. (1993). Infant macaque monkeys respond to pictorial depth. *Psychological Science, 4,* 93–98.

Mitchell, D. E., Kaye, M. G., & Timney, B. (1979). Assessment of depth perception in cats. *Perception, 8,* 389–396.

Timney, B., & Keil, K. (1999). Local and global stereopsis in the horse. *Vision Research, 39,* 1861–1867.

Animal Eye Movements

The brain receives visual information about the environment through photoreceptors in the retina, which convert light into electrochemical signals and ultimately into neural activity. This entry describes various experiments on animal eye movements. The spatial orientation of the retina with respect to objects in the world determines what kind of visual information is perceived. Therefore, in all mobile animals the position of the retina with respect to visual objects of interest must be controlled, to ensure that they are aligned and that the retina is stationary with respect to such objects long enough for visual inspection.

An interesting exception are heteropod mollusks, which have eyes that are shaped as a long narrow stripe only three to six receptors wide and several hundred receptors long. Their visual field of view is only a few degrees high and 80 to 180 degrees long. At any given time, therefore, most of the surroundings are not seen by the animal. To solve this problem, the heteropods systematically scan the environment with smooth rotating eye movements. It seems likely that the function of these movements is to enable the carnivorous mollusks to detect stationary objects in the surrounding water. In contrast, the eye movements of insects, higher crustaceans, cephalopods, and vertebrates have possibly evolved to detect motion, either self-generated as in locomotion, or caused by prey or predators, and that the reason for the maintenance of eye stability is the need to avoid confusion between eye movement and image movement.

The alignment of retina and visual object is controlled by neural mechanisms that either stabilize or shift gaze. Normally, this is achieved by moving the entire eye relative to the body; in some animals, however, only part of the eye is moved. Jumping spiders have excellent vision, with among the highest acuities in invertebrates. Two of their eight eyes, the anterior median eyes (AME), provide high visual acuity but small field of view, and the remaining six eyes provide lower resolution but broad field of view. The AME are long and tubular, which helps their resolution (longer focal length, more magnification) but which means they have a narrow field of view. The AME's have a narrow field of view, so the spider needs to point them in different directions to see objects at different locations. However, the spider cannot move the whole ocular mechanism because the lenses of the eyes are actually built into the carapace. Instead, a special set of muscles moves the retina around, and the lens stays fixed.

Gaze stabilization is also important for animals that cannot move their eyes in their heads at all. For example, insects need to stabilize their gaze during flight. They achieve this by performing very fast (2000 degrees/second) turns of their thorax in mid-flight. These thorax movements start first, and are followed by head movements that are later and more rapid to minimize the time of gaze shift and to maximize the time available for analyzing the surroundings.

Vertebrates use two mechanisms to stabilize the position of the retina with respect to visual objects: the *vestibulo-ocular reflex* and the *optokinetic nystagmus*. The vestibulo-ocular reflex (VOR) compensates for movements of the head and relies on vestibular input from the labyrinthine semicircular canals that respond to acceleration of the head. This information is carried by neurons within the vestibular ganglion and relayed to neurons in the vestibular nucleus. These neurons integrate the vestibular input and generate an appropriate compensatory motor response by innervating oculomotor neurons. The vestibular information is much faster than visual information. Thus, the compensatory eye movements generated by the VOR have a latency of 15 milliseconds (ms), whereas eye movements that rely on visual information typically have latencies of more than 70 ms. However, because the VOR relies on information about acceleration, it becomes increasingly unreliable during sustained head rotations with constant velocity and slow acceleration. In this situation, gaze stabilization relies on visual input, using the optokinetic nystagmus (OKN). The compensatory movements generated by this system are slow, but include the full range of motion of the eyes in the head. Long, sustained head movements can lead to large shifts of the eyes. When the eyes reach the outer corners of the oculomotor range, a corrective rapid eye movement moves the eyes in the same direction as the head rotation, which enables inspection of the oncoming visual scene. These rapid eye movements are generated by the same brainstem neurons that are responsible for saccades, and this system is likely the evolutionary precursor of the saccadic gaze-shifting system.

The evolution of gaze-shifting mechanisms was made necessary by specializations in the structure of the retina. In primates, the retina contains the fovea, a region of maximal photoreceptor density, which allows sampling of visual information with the highest spatial resolution. However, the fovea is limited to 1 to 2 degrees of visual space. Therefore, this zone of higher acuity must be directed to interesting parts of the visual field, which is being achieved by *saccadic eye movements*. The eye movements are rapid, and accurate shifts of the eyes with peak velocities of as much as 500 degrees are generated by a network of neurons in the brainstem.

Lateral-eyed species (such as rodents and ungulates) turn their heads in a series of rapid steps to scan the environment, but frontal-eyed species can turn their head smoothly through an angle of 180 degrees to fixate a target. In addition, they have usually developed the capacity to uncouple saccadic eye and head movements by suppressing the VOR. The higher-order gaze shifting commands that specify where and when the eyes should be shifted are generated in the superior colliculus (SC) and in a number of cortical areas, in particular the frontal eye field (FEF). Once the fovea is aligned with the object of interest, it has to be held on the object long enough that light from this region can be analyzed by the visual system. In the case of a stationary target, this is achieved by active *fixation*. For tracking a smoothly moving object, saccades are of limited use because the image soon slides off after being captured by the saccade and requires another saccade. In between the saccade cycles, there would be a loss of visual acuity. Therefore, the *pursuit* system evolved, which generates smooth tracking movements of the eyes that closely match the pace of the target. However, the visual system responds with a latency of 60 to 100 ms, which introduces a corresponding input lag into the neural feedback mechanism. The pursuit system overcomes this problem by using a feedforward mechanism to anticipate predictable motions of the target. During saccades and smooth pursuit, each eye usually moves in the same direction (these are called version movements), but in animals with binocular vision, for example primates, gaze shifts can also involve *vergence movements,* in which each eye moves in opposite directions. In this way, targets at different depth can be fixated upon.

Although gaze stabilizing mechanisms are automatic reflexes, gaze-shifting mechanisms can be initiated both by automatic and by voluntary, goal-directed processes. This voluntary control depends on a number of areas in the frontal and parietal cortex. These cortical eye fields are connected with the superior colliculus and the brainstem centers for the generation of saccades and pursuit movements both directly and indirectly through the basal ganglia. In recent years, this connection has been used to investigate cognitive processes, such as attention and decision making, by studying higher-order oculomotor areas.

Veit Stuphorn

See also Decision Making, Perceptual; Depth Perception in Pictures/Film; Eye Movements: Behavioral; Eye Movements: Physiological

Further Readings

Hikosaka, O. (2007). Basal ganglia mechanisms of reward-oriented eye movement. *Annals of the New York Academy of Sciences, 1104,* 229–249.

Leigh, R. J., & Zee, D. S. (2006). *The neurology of eye movements* (4th ed.). Contemporary Neurology Series. Oxford, UK: Oxford University Press.

Schall, J. D., & Thompson, K. G. (1999). Neural selection and control of visually guided eye movements. *Annual Review of Neuroscience, 22,* 241–259.

Scudder, C. A., Kaneko, C. S., & Fuchs, A. F. (2002). The brainstem burst generator for saccadic eye movements: a modern synthesis. *Experimental Brain Research, 142*(4), 439–462.

ANIMAL EYES

The structural organization of animal eyes provides the first step toward visual perception. There is substantial diversity among animal eyes, yet they tend to fall into a few distinct categories such as single chamber and compound eyes. Underlying nearly all light perception are visual pigments based on a family of proteins called opsins. These molecules are tightly linked to retinal, a chromophore related to vitamin A, which undergoes a conformational change in response to light exposure. This activates a photoreceptive cell (either by hyperpolarization, as generally is the case for vertebrates, or by depolarization, which is typical for most invertebrates) and, ultimately, after many computational steps, results in the animal's perception of light. Differences in eye design influence factors such as the direction from which a receptor cell absorbs light, or whether light is pooled from many directions. Therefore, the eye structure largely defines how specific animal eyes function, which frequently is near the physical limits of light. This entry describes how the eyes of different animals function.

From a Light-Sensitive Surface to a Camera-Type Eye

Notably, the most important features in eye design are aimed at gaining spatial resolution. Such resolution can be achieved by curving the photosensitive surface. In a flat surface, photoreactive cells absorb photons from different directions in space equally well. This is the case for light-sensitive spots of a limpet. In this animal, only light from beneath is shielded through the presence of screening pigments. If the surface curves inward (concave), as for example is the case for the slit shell mollusk *Pleurotomaria,* individual cells start to become selectively more activated by light from specific directions. This is referred to as spatial resolution, and this property tends to improve with increased curvature and a narrowing of the rim area of the eye pit. In photographic terms, the narrowing of the rim corresponds to reduction of the aperture. The resulting eye type, referred to as a pinhole eye, functions on the same principles as a pinhole camera and is exemplified in the cephalopod mollusk *Nautilus.*

Although pinhole eyes can yield good spatial resolution, their small aperture means that they perform poorly at low light levels. To increase

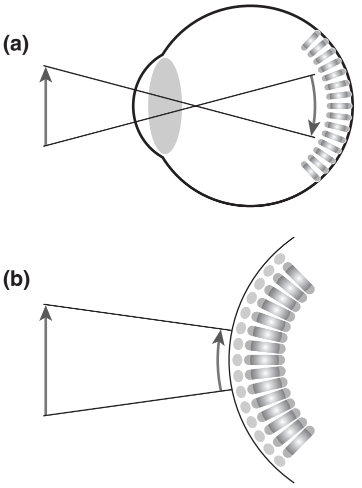


Figure 1 Schematic Organization of a Single-Chamber Eye (a) and a Compound Eye (b)

light capture, the pinhole or aperture has to be widened, which consequently results in a dramatic loss of spatial resolution. The solution in photography and animal eye evolution alike is the introduction of a lens, which allows capture of much more light while using the focusing power of the lens to maintain good resolution. The resulting "camera" or "single chamber" eye (Figure 1a) is found in vertebrates. Although most vertebrate eyes look fairly similar, organizational details do vary with species. For example, although mammals focus images by changing the shape (and hence the refractive power) of the lens, fish and certain amphibians adjust the focus by changing the location of the lens. The organization of the retina also differs in animals, often because of the animal's ecology or environment. For example, for us, the frontal visual space is particularly important, and we humans accordingly have a single fovea (area of increased visual acuity) within our binocular visual fields. Certain birds of prey need to simultaneously keep track of visual stimuli in the front and on the side. Accordingly, they have a second fovea within the monocular visual field of each eye. Another example is that animals living in flat environments, whether fish, mammals, or birds, tend to have horizontally oriented high-acuity areas, which allow them to see particularly well along the horizon.

Figure 2 Eyes of a Male Black Fly

Note: The compound eyes of a male black fly dominate the head and are divided into dorsal and ventral regions that together acquire visual input from nearly all directions in space.

Interestingly, the single chamber eye has also evolved multiple times in invertebrates. The most notable examples are the eyes of squid and octopi, which show remarkable similarities in eye design to those of vertebrate eyes. In addition, single chamber eyes with noteworthy resolving power are found among spiders, in the accessory eyes of certain adult insects, and as the primary eyes of some larval forms of insects that undergo complete metamorphosis (such as beetles and butterflies).

The Compound Eye

Most crustaceans and insects have evolved a different solution toward achieving spatial resolution: rather than curving the photoreceptive epithelium inward, evolution has led to a convex organization (Figure 1b) in these creatures. This outward curvature results in a spatially dependent activation of specific receptor cells that is similar to that of the concave epithelium, except that with the convex surface, the configuration does not allow for the presence of a single lens. Instead, each portion of the insect compound eye (referred to as ommatidium) is usually equipped with its own lens. Although externally the compound eye looks different from a single chamber eye, it essentially represents an inside-out version of our eye, and many organizational principles of the convex eye revolve around the same physical compromises made in the concave eye. For example, both eye types are characterized by trade-offs between how many points in space are resolvable and how much light each point captures. Although they are different in physical design, both eye types have actually converged on the same solution, namely, to devote only a portion of the eye to high-acuity vision.

As might be expected, compound eye design imposes major functional constraints. For example, because each lens tends to be small, these eyes are more likely to operate near the diffraction limit of visible light. For that reason, a compound eye would have to be substantially larger (about one meter in diameter) to achieve levels of spatial resolution that are comparable with those of human eyes. One advantage of compound eyes is that they are well designed to seamlessly monitor many directions in space. In fact, it is quite common to find insects with nearly 360° vision (Figure 2).

Elke K. Buschbeck

See also Animal Color Vision; Eye: Structure and Optics; Eyes: Evolution of; Vision; Visual System: Evolution of

Further Readings

Buschbeck, E. K., & Friedrich, M. (2008). Evolution of insect eyes: Tales of ancient heritage, deconstruction, reconstruction, remodeling, and recycling. *Evolution: Education and Outreach, 1,* 448–462.

Land, M. F., & Nilsson, D. E. (2002). *Animal eyes.* Oxford, UK: Oxford University Press.

Nilsson, D. E., & Pelger, S. (1994). A pessimistic estimate of the time required for an eye to evolve. *Proceedings. Biological sciences/The Royal Society, 256*(1345), 53–58.

Tsonis, P. A. (2008). *Animal models in eye research.* San Diego, CA: Academic Press.

Warrant, E., & Nilsson, D. E. (2006). *Invertebrate vision.* Cambridge, UK: Cambridge University Press.

ANIMAL FREQUENCY AND PITCH PERCEPTION

Pitch is a perception that is evoked by simple tones and many different types of complex sounds, including speech and music. *Pitch* refers to the subjective experience produced by a sound. The pitch of a sound is closely related to the frequency of a pure tone or to the fundamental frequency of complex sounds, but does not refer to the specific frequency content of the sound. As such, pitch is what results from how the brain analyzes and processes the frequencies present in a sound. This entry describes sensitivity ranges of different animals, the ability to detect frequency differences, and animal perception of naturally occurring sounds.

Animals most likely experience sensations that are similar in many respects to the pitch percepts experienced by human listeners, but scientists cannot directly ask animals about the subjective experience in a way that can be asked of human listeners. To understand animal perceptions, scientists must ask the animal indirectly how it perceives the sound. This is done by training or conditioning the animal to make a behavioral response to a particular feature of a sound that will then be varied. For pitch perception, the frequency content of a sound can be varied. When the animal makes a correct behavioral response, it usually receives some sort of feedback, such as a food reward or avoiding a mild electric shock. Although scientists have no way of knowing exactly what the animal's subjective experience is, from the behavioral responses of the animals, scientists can determine what features of the sounds the animal is analyzing and thus make inferences about what the animal experiences.

It is not surprising that the way in which the brains of animals process the frequency content of sounds is similar, in many respects, to that of the human brain because the general structure of the auditory brain is similar in animals. A major difference among animals is the range of frequencies that animals are able to perceive, and these differences are related to the structures that constitute the outer, middle, and inner ears. The hearing range of frequencies for most fish is fairly low, between 50 and 500 hertz (Hz). However, some fish have accessory auditory structures such as swimbladders and tiny bones that connect the swimbladder to the inner ear. These structures act like the middle ear in land animals and increase the sensitivity to higher frequencies to about 2,000 Hz. Amphibians and reptiles typically have a range of hearing between 100 and 2,000 Hz. Most birds have a relatively narrow hearing range with best sensitivity between 1,000 and 3,000 Hz, but some species of owls have excellent sensitivity as high as 10,000 Hz. Unlike birds, the range of hearing is quite heterogeneous in mammals. For example, the range of hearing for humans is 20 to 20,000 Hz, whereas elephants can hear frequencies below 20 Hz with great sensitivity, and dolphins and bats can hear frequencies above 100,000 Hz with great sensitivity. The range of hearing will determine how the animal perceives the outside world.

Another aspect of frequency processing is the ability of animals to detect differences in the frequency of a simple tone. All animals tested can detect changes in the frequencies of simple tones, but it is well established that human listeners can detect smaller frequency differences in a tone than can all other animals tested. That is, animals are less sensitive to frequency changes than are humans. These findings indicate that animals have a psychological dimension (i.e., a percept) that corresponds to the frequency of a tone. Behavioral data obtained using simple pure tones can be directly related to the map of frequency-to-place that exists along the basilar membrane in the cochleae of birds and mammals. That is, for simple tones, the frequency map along the basilar membrane does correspond to the pitch map.

Most naturally occurring sounds are made of multiple frequencies, and when the frequencies are harmonically related, a single pitch at the fundamental frequency is evoked in human listeners. For example, a sound made of five tones at 1,200, 1,400, 1,600, 1,800, and 2,000 Hz will evoke a pitch that corresponds to 200 Hz. Note that the sound itself contains no energy at the fundamental frequency of 200 Hz, but nevertheless, there is a strong pitch at 200 Hz. The phenomenon of having a perception of a pitch from a sound containing no energy at the fundamental frequency is known as the *pitch of the missing fundamental.* Like humans, birds and mammals respond to these same types of sounds as if they perceive the pitch of the sound to be at the missing fundamental frequency. The auditory areas of the cerebrum appear to be important for the ability of animals and humans to detect changes in the pitch of the missing fundamental. Nevertheless, animals require a larger difference in fundamental frequency than do humans in their ability to detect that the fundamental has changed. Moreover, different types of complex sounds having different frequency contents can evoke the same pitch percept, but the frequency map along the basilar membrane does not correspond to the pitch map for complex sounds.

Changes in the pitch of the fundamental frequency of complex sounds give rise to melody perception, and the pattern of pitch changes between the individual notes of the melody is called the *frequency contour*. Although birds and mammals can discriminate frequency contours, it appears that animals base their discrimination primarily on the absolute frequencies of the tones, whereas humans base their discrimination primarily on the intervals between individual frequencies. Consequently, human listeners can recognize a melody regardless of whether the individual notes change in the same register or whether the individual notes change in octave registers. This phenomenon is called *octave generalization,* and it has been demonstrated to also occur in monkeys. However, in general, octave generalization does not appear to be as salient in birds and mammals as it is in human listeners.

William P. Shofner

See also Audition: Pitch Perception; Auditory Frequency Selectivity; Auditory Processing, Central; Auditory Processing, Peripheral; Music Cognition and Perception

Further Readings

Fay, R. R. (1988). Comparative psychophysics. *Hearing Research, 34,* 295–306.

Heffner, H. E., & Heffner, R. S. (2007). Hearing ranges of laboratory animals. *Journal of the American Association for Laboratory Animal Science, 46,* 11–13.

Plack, C. J., Oxenham, A. J., Fay, R. R., & Popper, A. N. (2005). *Pitch: Neural coding and perception.* New York: Springer Science.

Tomlinson, R. W. W., & Schwartz, D. W. F. (1988). Perception of the missing fundamental in non-human primates. *Journal of the Acoustical Society of America, 55,* 1061–1069.

Wright, A. A., Rivera, J. J., Hulse, S. H., Shyan, M., & Neiworth, J. J. (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology: General, 129,* 291–307.

ANIMAL MOTION PERCEPTION

Most animals move around in their environment, and through evolutionary time, this has resulted in specializations in their sensory systems and brains, to detect motion, as described in this entry. These patterns of movement can occur as moving images in vision, as moving sounds in hearing, and as moving tactile patterns in touch. Visual motion can be simply divided into two classes; the first, called *object motion,* is produced when some object moves relative to stationary objects in the world. The second class of motion, called *self-motion,* is produced by the observing animal itself moving its body, head, or eyes and thereby creating movement of the entire image across its visual field.

Although visual object motion can be produced by nonliving things such as rain, snow, leaves, waves, or clouds, most visual object motion is produced by the movement of other animals. It is advantageous for an animal to notice the movement of other animals because other animals may be the viewing animal's prey, a predator, or one of its own species and detecting them will greatly help the animal capture prey, avoid predators, and engage in social interactions, respectively. The other distinctive class of motion patterns is produced by an animal itself moving through a world of stationary objects, and these visual "flow patterns" are used to control its own movements, steer through space, obtain depth information through motion parallax, avoid obstacles, and even help maintain its posture and balance. Not surprisingly, there are separate areas of the visual system to process *moving objects* and *self-produced motion* patterns, which in turn generate distinctive behavioral responses, and they are found in both invertebrates and vertebrates alike.

Visual Neurons

Early research on frogs and insects showed that some visual neurons were specialized to respond only to relatively small moving visual stimuli, and often these preferred a particular direction of motion. These directional specific neurons were sequence detectors. If two adjacent areas in the visual field (say A and B) were stimulated by a moving stimulus that moved over A then B, the neuron would fire to signal this event. However, if the direction were reversed so that B then A was stimulated, the neuron would not fire, but would be inhibited. Early experiments showed that two flashed stimuli one at A and then the next at B after a short delay, also stimulated these movement neurons, but again, they did not fire when B then A were flashed without movement. Thus, these neurons respond to *apparent motion* as well as real motion. These motion-detecting neurons that respond to small moving things are precisely the stimuli that capture an animal's attention and typically produce a turning of the head and eyes ("orienting response") to place the object's image onto the fovea or retinal area of high resolution. In frogs, these were nicknamed "Bug Detectors" because frogs turn and snap at bug-like moving patterns. Typically, larger moving stimuli produce escape or avoidance behavior in animals.

Movement and direction-specific neurons have been found in the visual system of all invertebrate and vertebrate animals where moving stimuli have been used to interrogate the visual system, and so it seems safe to assume these neural specializations are universal. Moreover, wherever motion-detecting neurons occur, it seems that the underlying mechanism in all species is produced by delayed asymmetric lateral inhibition. Object motion sensitive neurons have been found in flies, bees, locusts, crickets, praying mantises, and dragonflies, and also in fish, salamanders, turtles, frogs, toads, chickens, pigeons, owls, mice, rats, cats, ferrets, wallabies, and several species of monkey. Perhaps the most extensive studies have been conducted on the middle temporal cortical area of the macaque monkey cortex.

Camouflage

Many animals freeze when they see something moving, and this behavior ensures that object motion detecting neurons in other observing animals will not be triggered. Together with their protective coloration, which often matches the features of the environment they inhabit, this makes animals difficult to see and for figure ground boundaries that define their shape to be resolved. However, should the background be moving, such as when the wind moves branches or leaves, the animals often move with the background. Walking stick insects rock back and forth as they advance and thus mimic twigs waving in the wind, and Australian horned devil lizards and some species of snake (e.g., oxybelis) use this strategy to foil the object motion detectors of their predators. Moreover, when a predatory animal such as members of the cat family, or jumping spiders stalk their prey, it seems they are attempting to keep their own speed below their prey's threshold for detecting motion, and their crouching posture ensures their legs, which always move with a faster velocity then their bodies, are not visible to their prey targets.

Biological Motion

In courtship displays of many animals, species-specific characteristic motion patterns are produced by the courting pair. In this case, one might assume that there are specific object motion detectors in the brains of these animals to recognize these patterns that play an important role in sexual selection.

When an animal walks, flies, or swims through its environment, the images of stationary objects produce specific patterns of motion across its entire visual field of view. These patterns of *optic flow* can also be produced by an animal's head or eye movements. These self-produced optic flow patterns do not activate object motion detecting neurons, but instead, other visual neurons in specialized parts of the brain of both vertebrates and invertebrates do respond best to large patterns of optic flow. This specialized pathway acts together with the vestibular system to help control an animal's balance and posture and to provide information about how the animal is moving through space. When human observers view an IMAX theater movie, this motion system is activated and produces the compelling illusion the observer is moving through space, which is termed *vection*. Animals placed in a rotating striped drum also must experience this because they make body head and eye movements in an attempt to keep the large moving image stable on their retinae.

Finally, many animals seem to be able to calculate the *time to collision* with an object that is directly approaching or looming toward them and likewise to compute their own time to collision with stationary objects. In the first case, this could be a predator attacking them and would allow them to take evasive action, and in the second case, it provides information so they can land on branches, fold their wings in a streamlined fashion if they are a diving bird, run thorough a thicket, and otherwise skillfully maneuver through their environment. Research has shown that insects landing from flight, hummingbirds, kestrels, and kingfishers hovering in the same spot, and gannets folding their wings while diving for fish are all computing time to collision from the expanding images of the surfaces they are approaching. Again, specialized cells have been found in the visual system of animals that are calculating this time to collision.

Barrie Frost

See also Camouflage; Ecological Approach; Motion Perception; Motion Perception: Physiological; Self-Motion Perception

Further Readings

Dittrich, W. H., & Lea, S. E. G. (2001). Motion discrimination and recognition. In R. G. Cook (Ed.), *Avian visual cognition.* Retrieved from http://www.pigeon.psy.tufts.edu/avc/dittrich

Masland, R. H., & Albright, T. (Eds.). (2008). Vision. In A. I. Basbaum (Series Ed.) & R. H. Masland & T. Albright (Vol. Eds.), *The senses: A comprehensive reference* (Vol. 2, Chapters 2.10 to 2.13). New York: Elsevier.

APHASIAS

Aphasia is a neurological impairment of the human language system. Language includes several subsystems (speech sounds, grammatical structure, meaning) and is conveyed and understood through various input and output channels (spoken, written, gestured). In aphasia, each of these channels is affected to some degree, with difficulty ranging from minimal to severe. Minimal deficits include being slower to understand reading material, to get a joke in conversation, or to choose a precise word. Severe aphasia impairs even simple language tasks, such as matching a single printed word to the object it represents, or providing one's name and address.

Aphasia results from damage that is nearly always isolated to the left side of the brain. This damage interrupts the blood supply to the brain and starves the affected areas of the oxygen and glucose they need to survive. After several minutes, those brain regions die.

Aphasia is often divided into subtypes. These subtypes are distinguished roughly by relative strengths and weaknesses in language channels and subsystems. Different aphasia types also have been linked with different locations of damage within the left brain. Robert Wertz and colleagues summarize the controversy about the value and validity of these subtype distinctions. Specifically, within each aphasia type, there is a lot of individual variation in the presentation of strengths and weaknesses, and in the location of brain damage. Aphasia also has been described as nonfluent versus fluent, which refers to the average length of uninterrupted stretches of spoken language. Broca's, transcortical motor, and global aphasia are all nonfluent aphasias with relatively short stretches of uninterrupted output, and Wernicke's, conduction, anomic, and transcortical sensory are fluent, with essentially normal phrase length. Some practitioners refer to aphasia as expressive or receptive. It has been argued that these descriptors are neither valid nor informative because people with aphasia have both expressive and receptive components to their disorder, to some degree. Aphasia also is sometimes described as anterior versus posterior. These adjectives are not widely used because they describe a general location in the brain rather than the language characteristics of aphasia.

Aphasia is not a deficit of basic sensory systems—hearing, vision, touch—or of motor function. However, aphasia overwhelmingly affects older adults, so concurrent problems with hearing or vision are common. This entry discusses perception in aphasia, implications of research on perception in aphasia, and mirror neurons and perception of actions in aphasia.

Perception in Aphasia

Speech Perception

Speech perception refers to a set of processes that decode and interpret speech sounds (consonants and vowels) and patterns of speech melody (intonation). The study of speech perception deficits in aphasia can be complicated by co-occurring hearing loss.

Some people with aphasia have trouble perceiving time changes in components of speech sounds and patterns. Aphasia caused by relatively posterior brain damage, such as Wernicke's or conduction aphasia, has been associated with difficulty detecting a brief silent interval in a noise signal, or identifying which tones in a series are closest together in time. Timing parameters like these are important for distinguishing among consonants but not among vowels, and indeed, aspects of consonant perception are more impaired in aphasia than is vowel perception.

Aphasia caused by more anterior brain damage, such as Broca's aphasia, may yield difficulty with aspects of speech frequency (roughly, melodic characteristics). Problems may occur in picking a tone out of a noise signal, or determining if two tones are the same. There may also be difficulty identifying the meaning of certain melodic contours of spoken language. For example, echo questions may be hard to identify. Echo questions are statements with the melodic contour of questions (e.g., "He's going to the store?").

Aspects of frequency perception have also been examined in aphasic speakers of tonal languages. Tonal languages are those, such as Mandarin Chinese or Thai, in which a word's meaning is wholly determined by its melodic pattern. Each tonal language has a fixed number of these melodic patterns, which are also called tones. Tonal language speakers with aphasia have difficulty identifying their language-specific tones, although they do better than people who do not speak tonal languages. There is not enough research on aphasic speakers of tonal languages to determine how this deficit influences the understanding of everyday communication.

For English speakers with aphasia, evidence indicates that problems with speech sound perception do not predict the ability to understand natural language in everyday situations. This may be largely because of the ability of many individuals with aphasia to use top-down perceptual processes to compensate for difficulties in processing speech sounds. Top-down processes help fill in gaps in sensory input. These gaps can occur when an input signal is obscured, incomplete, or otherwise difficult to detect, such as when there is noise in the environment or when unfamiliar words are used. Top-down processes in speech perception rely on sources of information beyond the speech signal itself, such as the perceiver's prior knowledge and expectations, and the remainder of the language utterance or situation.

Music Perception

Music perception in aphasia has rarely been explored. Anirudh Patel reasons that music and language have some common grammatical features. Specifically, both consist of elements (chords, words) that are rapidly combined in sequence into (more and less) acceptable structures. Patel's recent work suggests that in aphasia, grammatical problems in music perception co-occur with grammatical problems in language comprehension. In one experiment, people with aphasia judged whether all chords in a sequence were a good fit musically. The sequences of most interest contained a chord that was out of place, from a harmonically distant musical key. Participants also judged the acceptability of spoken sentences, including some with errors in subject-verb number agreement (e.g., "The *sailors* call for the captain and *demands* a fine bottle of rum"). On both tasks, participants with aphasia were less accurate than were participants without aphasia. Performance of both groups on the music task predicted performance on the language task.

Visual Perception

Visual perception is another rare target of investigation in aphasia. At a global level, visual perception in aphasia is reasonably spared. Even people with relatively severe aphasia can match like objects with each other, and often like pictures. Severely impaired individuals can also copy alphabet letters, simple geometric shapes such as squares and triangles, and often their own names. However, detailed assessment of visual perception in aphasia is nearly nonexistent. One study reports a problem picking an image out of a background when damage was relatively posterior in the brain. Another study indicates that people with aphasia have difficulty determining whether mouth movements are the same or different. This same discrimination difficulty is also evident when only the sounds that correspond with the mouth movements are presented.

Perception of Space

Little is known about spatial perception in aphasia. Some people with aphasia do better on a wide range of language tasks—naming pictures of objects, reading words, generating nouns, determining synonyms, understanding sentences, telling a fairy tale—when their bodies are oriented to the left side of space than when oriented to the right side. People without brain damage perform these language tasks without error. Note here that damage to one side of the brain affects the opposite side of the body, so language performance in aphasia is enhanced when the body is oriented toward its good side. Other evidence suggests a potential spatial influence on sentence production and comprehension in aphasia. A man with grammatical language deficits was shown drawings of two stick figures, one of whom was doing something to the other, such as pushing. This man consistently chose the figure on the left as the one performing the action.

Odor Perception

Odor perception has rarely been explored in aphasia. People with aphasia are reportedly able to discriminate between common aromas, such as coffee, fish, and soap. However, they may have difficulty matching common odors to pictures or objects. Although results are mixed, the largest available study hints that odor may facilitate naming ability in aphasia. Eleven of 30 participants did better naming an object when provided its odor than when presented a picture, and more severely impaired individuals benefited the most from odor cues.

Implications of Research on Perception in Aphasia

Investigation of perception in aphasia has many potential implications. First, if the available evidence can be replicated, there will be new tools to investigate for their effectiveness in language treatment, including musical, aromatic, and spatial orientation cues. As an example, presenting a to-be-named item together with its odor in aphasia treatment is a way to leverage perceptual strengths to improve language performance, in this case naming, for some individuals with severely reduced naming abilities. Similarly, orienting an aphasic individual's body toward the good side of space during treatment may enhance performance on a range of language tasks. Treatment that uses areas of strength to scaffold areas of weakness may contribute to establishing and reinforcing new brain pathways that subsequently help the person with aphasia communicate more effectively in those areas of weakness.

In a related vein are implications for brain plasticity research, or the study of the brain's ability to reorganize after it is damaged. Brain plasticity research aims to identify brain regions that can help support a function, even if they normally do not. Most brain plasticity research in aphasia examines how brain regions are enlisted to spur improvements in the subsystems and channels of language. A few studies have addressed basic timing and frequency parameters of speech perception. In one example, when aphasia damaged a left-brain region that normally supports speech perception, new brain activity was evident in the same region on the right side of the brain.

Investigation of perception in aphasia also has the potential to enrich theories of the relationships among brain function, language ability, and perceptual phenomena and subsystems. For example, research on aphasia and perception could help refine proposals about what parts of the brain are necessary for various perceptual processes. When brain damage disrupts a particular function, such as an aspect of language or perception, one can infer that the brain regions that are damaged are typically necessary to support that function.

Mirror Neurons and Perception of Actions in Aphasia

Mirror neurons are specific brain cells in monkeys that are engaged in two conditions: when the monkey watches an action being performed and when the monkey itself performs the action. Mirror neurons have not been identified in humans, and they were originally described to link basic observational and motor skills. However, research findings in humans, and for higher-level skills such as language, are increasingly being interpreted through the prism of mirror neurons. In aphasia, for example, deficits in interpreting pantomimed actions and in comprehending printed action words have been linked by some to portions of a human analogue of the mirror neuron system. Among others, Luca Turella and his colleagues have recently outlined the as yet highly controversial nature of such interpretations.

Connie A. Tompkins

See also Cortical Reorganization Following Damage; Language; Mirror Neurons; Music Cognition and Perception; Neuropsychology of Perception; Speech Perception

Further Readings

Blumstein, S. E. (2004). Phonology and adult aphasia. In R. Kent (Ed.), *The MIT encyclopedia of communication disorders* (pp. 366–369). Cambridge: MIT Press.

Coslett, H. B. (1999). Spatial influences on motor and language function. *Neuropsychologia, 37,* 695–706.

Hickok, G. (2000). Speech perception, conduction aphasia, and the functional neuroanatomy of language. In Y. Grodzinsky, L. Shapiro, & D. Swinney (Eds.), *Language and the brain: Representation and processing* (pp. 87–104). San Diego, CA: Academic Press.

Mummery, C. J., Scott, S. K., Ashburner, J., & Wise, R. J. S. (1999). Functional neuroimaging of speech perception in six normal and two aphasic patients. *Journal of the Acoustical Society of America, 106*(1), 449–457.

Patel, A. D., Iversen, J. R., Wassenaar, M., & Hagoort, P. (2008). Musical syntactic processing in agrammatic Broca's aphasia. *Aphasiology, 22*(7–8), 776–789.

Schmid, G., & Ziegler, W. (2006). Audio-visual matching of speech and non-speech oral gestures in patients with aphasia and apraxia of speech. *Neuropsychologia, 44*(4), 546–555.

Smithpeter, J. V. (1976). A clinical study of responses to olfactory stimuli in aphasic adults. Retrieved May 8, 2008, from Clinical Aphasiology archive: http://aphasiology.pitt.edu/archive/00000038/

Turella, L., Pierno, A. C., Tubaldi, F., & Castiello, U. (2009). Mirror neurons in humans: Consisting or confounding evidence. *Brain & Language, 108*(1), 10–21.

Wertz, R. T., Dronkers, N. F., & Ogar, J. (2004). Aphasia: The classical syndromes. In R. Kent (Ed.), *The MIT encyclopedia of communication disorders* (pp. 249–252). Cambridge: MIT Press.

AROMATHERAPY

Aromatherapy is the use of essential oils for the treatment of physical and psychological health. An essential oil is the distilled essence of a particular plant typically combined with a vegetable oil base. Contemporary aromatherapy is based on the inhalation of aromas rather than their ingestion or application. Thus, odor sensation and perception are responsible for the reported effects. This entry describes the history of aromatherapy, aromachology, and theories of aromatherapy.

History

The concept of aromatherapy came from ancient cultures that believed in the healing properties of plants. Many plants have therapeutic properties, and the basic derivatives of many modern medicines are plant-based. The term *aromatherapy* (originally *aromatherapie*) was coined in the late 1920s by the French chemist Rene-Maurice Gattefosse, who began the exploration of essential oils for their healing powers after an explosion in his laboratory left his hand badly burned and the accidental soaking of his injury in pure lavender oil produced a miraculously rapid healing.

Popular aromatherapy is primarily rooted in folkloric traditions rather than scientific research and a wide range of therapeutic effects are often attributed to any one aromatic plant. Here are some examples of the purported effects for several commonly "prescribed" essential oils: Sandalwood aroma is sedating, relaxing, and beneficial for treating anxiety, depression, and insomnia. Rosemary clears the mind and stimulates memory. Lavender is uplifting, soothing, and helpful for reducing stress, anxiety, depression, and insomnia. Indeed, authors of aromatherapy textbooks make a number of extraordinary and sometimes contradictory claims. For example, juniper oil is said to have 17 different properties ranging from aphrodisiac to sedative. Aromatherapy today is usually practiced in conjunction with massage, and one of the major claims of aromatherapy is a reduction in anxiety. Considerable evidence shows that massage alone reduces anxiety. If massage is effective, then aromatherapy plus massage is also effective. So "aromatherapy works, even if it doesn't," as several reviews of the literature have noted, at least for the reduction of anxiety. Thus, the wide range of therapeutic effects attributed to a plant's scent alone appear to be unfounded.

Aromachology

The term *aromachology* was coined by the Sense of Smell Institute in 1982, to distinguish between nonscientific claims for odors (aromatherapy) and the scientific analysis of olfactory effects on mood, physiology, and behavior (aromachology). Aromachology research and evidence must meet the following empirical criteria: (a) theory guided goals and clear hypothesis testing, (b) fragrances assessed using appropriate experimental methodology, (c) sufficient and representative subject populations and appropriate contrasting control groups, (d) data analysis with suitable statistical methods, and (e) results vetted by scientific peers and accepted for publication in reputable journals.

A rigorous analysis of the data obtained from aromachology research has shown that various specific odors can significantly alter mood, cognition, physiology, and behavior. However, inconsistencies in the chemical nature of the odors used, the dependent measures assessed, and the findings obtained are observed. For example, in one experiment, subjective emotional changes in the presence of lavender odor were found but no physiological effects were observed, whereas in another test of lavender, physiological changes were shown but there were no concomitant changes in self-reported mood.

Theories

A pharmacological hypothesis and a psychological hypothesis have been proposed to explain how and why odors produce the effects that have been observed on various aspects of mood, physiology, and behavior.

The pharmacological hypothesis proposes that the effects of aromas result from an odor's direct and intrinsic ability to interact and affect the autonomic nervous system/central nervous system or endocrine systems. In support of this proposition, lavender has been shown to act postsynaptically (after synaptic transmission between neurons) where it is suggested to modulate the activity of cyclic adenosine monophosphate (cAMP). A reduction in cAMP activity is associated with sedation. Linalool, a principal component of lavender essential oil, has also been found to inhibit glutamate binding, which may have sedative effects. It is therefore possible that lavender produces sedative effects via these neuropharmacological mechanisms.

The psychological hypothesis proposes that odors exert their effects through emotional-associative-learning, conscious perception, and belief-expectation. Emotional-associative-learning with odors is seen when an odor that has been associated to an emotional state is able to immediately elicit those emotions when later smelled. The reason for the instant emotional response is the uniquely direct and intimate neuroanatomical connection between the olfactory cortex and the amygdala, the region of the limbic system where emotional responses are processed. The physiological responses that then may be observed are simply the physiological sequelae of the emotional state elicited by the odor. Under associative learning, idiosyncratic associations will alter the way an aroma is perceived and what effects it may produce. Illustrating this, one study testing the influence of aromas on the experience of pain found that only odors that participants *self-selected* as pleasant were able to have a positive effect.

Expectation and belief have also been shown to play a major role in aromachology outcomes. In several studies, a water control condition (no scent) was able to produce the expected behavioral and physiological changes in accord with what the participants *believed* about the scent that was present.

Aromatherapy proponents make much of the superior properties of natural versus synthetic chemicals, yet this has not been empirically addressed. The chemical nature of the specific odors tested in aromachology research has varied widely across laboratories, including the use of synthetic, natural or different chemical versions (e.g., isomers) of a particular odor. A number of individual difference factors also mitigate and modulate the responses elicited by an odor, including culture, experience, gender, and personality. For example, in studies examining the effects of odors on sleep, men and women were found to be differently affected by lavender and peppermint, respectively.

At present, there is better empirical support for the psychological hypothesis and that the *perceived* quality of the odor is the most relevant factor for determining how an individual responds to it both emotionally and physiologically. That is, perceptual effects rather than intrinsic chemical properties best explain how specific scents can influence mental and physical states. Nevertheless, pharmacological mechanisms might also be operative with at least some compounds such as (–) linalool, a key aromatic compound in lavender.

Rachel S. Herz

See also Context Effects in Perception; Fragrances and Perfume; Mind and Body; Olfaction

Further Readings

Campenni, C. E., Crawley, E. J., & Meier, M. E. (2004). Role of suggestion in odor-induced mood change. *Psychological Reports, 94,* 1127–1136.

Chen, D., & Dalton, P. (2005). The effect of emotion and personality on olfactory perception. *Chemical Senses, 30,* 345–351.

Cooke, B., & Ernst, E. (2000). Aromatherapy: A systematic review. *British Journal of General Practice, 50,* 493–496.

Herz, R. S. (2009). Aromatherapy facts and fictions: A scientific analysis of olfactory effects on mood, physiology and behavior. *International Journal of Neuroscience, 119*(2), 263–290.

Lis-Balchin, M., & Hart, S. (1999). Studies on the mode of action of the essential oil of lavender (Lavandula angusifolia P. Miller). *Phytotherapy Research, 13,* 540–542.

Vickers, A (2000). Why aromatherapy works (even if it doesn't) and why we need less research. *British Journal of General Practice, 50(455),* 444–445.

ART AND PERCEPTION

See Pictorial Depiction and Perception

ASSISTIVE TECHNOLOGIES FOR THE BLIND

Assistive technologies are devices or software that assist the accomplishment of tasks by a blind or visually impaired person, usually providing feedback via sound, touch, or optimal use of remaining vision. Technology is only one tool in the rehabilitation process, but for some tasks it can be an important one. This entry describes the range of assistive technologies that are available for the blind.

Technology for Reading, Writing, and Note Taking

For straightforward text, reading machines and optical character recognition software (available in both desktop and handheld units) can scan printed material and read it aloud in artificial ("synthetic") speech. Audio recordings from the National Library Service and other sources provide access to many publications. These "talking books" are now available on CDs and other digital media. The National Federation of the Blind provides a service giving access to spoken versions of newspapers and magazines, downloaded via telephone, with sections and articles selected via touch-tone commands.

Braille, a tactile code using raised dots for letters and numbers, is used for reading and writing by blind individuals. A simple stylus is often used to punch raised dots into paper for short notes and labels, but Braille translation software and printers ("embossers") for computer output are now commonplace. Mechanical and electronic Braillers are also available, with a Braille keyboard (six keys corresponding to the dots in a Braille character, and a space bar) coupled to a print head that embosses the Braille characters on paper. Electronic "refreshable" Braille displays mechanically present a line of Braille characters that are replaced by the next line when the reader activates a control switch. These are used for displaying computer output and for compact personal Braille notetakers, which provide much of the functionality that a notebook computer does for the sighted.

For an individual with low vision, some books are available in large print, and individuals can readily produce their own large print text using a computer. The inverse of the Snellen acuity fraction approximates the magnification needed to read small print. For example, someone with 20/60 acuity is likely to need 60/20 = 3 times magnification for normal reading. For this amount of magnification, common optical magnifiers (available with built-in illumination) will suffice. "Reading telescopes" attached to spectacle frames facilitate comfortable posture, farther away from the reading material. A strong positive spectacle lens or "add" focused an inch or two away from the face gives 8 to 16 times magnification compared with a normal reading distance of 16 inches. Even larger magnifications (as much as about 60) can be obtained with closed-circuit television (CCTV) magnifiers that display a magnified video image on a television-type monitor. (At high magnifications, only two or three letters may be visible at a time, slowing reading and making scanning more difficult). Writing materials can also be placed under the unit's camera. The advent of liquid crystal displays (LCDs) has produced handheld versions with a miniature camera underneath an inch-thick screen measuring about 6 by 3 inches, providing as much as 10 to 12 times magnification with a larger field of view than an optical magnifier. Most electronic magnifiers offer a contrast reversal option (i.e., displaying white print on a dark background), which can reduce glare. Some can be mounted on the head, or used for reading text at both near and far (e.g., the blackboard in a classroom).

Appropriate lighting is extremely important in low vision; performance in poor light is often much more affected than standard acuity measurements would indicate. A strong reading light positioned over the shoulder and focused on the reading material to reduce glare can improve the situation radically. Special reading lights are often dispensed by low vision clinics and catalogs.

Access to Computers and the Internet

Nonvisual access to the information on a computer screen is obtained using a screen reader software package. This software allows the user to select text displayed on any part of the screen, and convert it into speech or Braille output. Keyboard commands are used to navigate around the screen. Screen readers allow access to most common office software applications (word processors, spread sheets, etc.) using text and numbers. They cannot interpret pictures, however, so access to pictorial or graphical information is not provided. Screen readers are also used to surf the Web. Because they cannot interpret pictures, a recent Web Access Initiative has issued recommendations for Web site authors to add descriptive text labels (known as "alt text") to each image.

The most common form of output for screen readers is synthetic (computer-generated) speech, which practiced blind users can use at several times the normal speed. Refreshable Braille displays are another output option, but are considerably more expensive.

For those with severely impaired vision, special screen magnification software can provide high magnification and full functionality with different applications. Modern computer operating systems and Web browsers come with some screen magnification options built in. Another often-ignored option for low-vision computer access is simply to buy a larger screen.

Access to Communications and Portable Computing Devices

The complex menus of modern cell phones create difficulties even for a normally sighted person—particularly an elderly one. Most phone networks offer one phone model that has speech output for at least basic functions. Speech-output cell phone "screen readers" have also emerged, similar in concept to computer screen readers. For low-vision users, special phone models with large buttons, numerals, and displays are available.

Most personal digital assistants (PDAs), pocket PCs, MP3 players, and other portable computing devices have small displays that are hard for a visually impaired person to see, and menu functions that are largely inaccessible without vision. Occasionally, an accessible version of a mainstream product appears, but special-purpose talking PDAs with Braille keyboards and screen reading software have been developed by assistive technology manufacturers.

Technology for Independent Travel

The most common technology used by blind travelers is the long cane—usually long enough to reach two steps ahead while striding forward. Tapping the cane tip produces sound echoes from nearby objects, such as walls, parked cars, and entryways. Users listen to these subtle echo patterns to glean additional information about the nearby environment—a technique known as echolocation.

Electronic travel aids (ETAs) developed to assist blind pedestrians include obstacle detection devices that detect echoes from ultrasonic or infrared signals emitted by the device as they reflect off environmental surfaces. This information may be displayed as audio or vibration. More sophisticated FM ultrasonic aids can produce an audible output signal whose pitch is proportional to range, and with a timbre that varies with the nature of the target.

Braille signs can assist orientation, but have to be found before they can be read. This problem led to the development of Remote Infrared Audible Signage (RIAS) to provide the blind user with the ability to locate and read signs remotely. Transmitters at sign or landmark locations transmit infrared signals that are detected and converted to speech by a receiver carried by the user, who can determine the sign direction by aiming the receiver for maximum clarity. Talking navigation aids using the global positioning system (GPS) to establish a blind user's location are also available. These cannot be used indoors or in underground transit, and accuracy is limited near tall buildings, but in more open outdoor areas, accuracy is excellent. Audible pedestrian signals can be installed at intersections to provide information on safe crossing times.

Blind individuals have historically had extremely limited access to street maps. New tactile map automated production (TMAP) technology now makes it possible to request a tactile street map centered on any address via the system Web site. The resulting map can be downloaded and printed on a Braille embosser. Audio map information is also available in conjunction with the various GPS systems mentioned above.

Environmental modifications such as painting the edges of steps and platforms in a contrasting color can greatly aid travel safety for individuals with low vision. Hand-held telescopes can help read street signs or items on a menu behind a counter. Spectacle-mounted manual and autofocus telescopes, and high-powered (up to 40×) hand-held electronic video zoom telescopes are also available. Night travel can be aided by strong flashlights or night vision telescopes or goggles.

Access to Graphical and Pictorial Information

For blind individuals, access to pictorial and graphical information is still extremely difficult. Verbal image description can be used in such tasks as describing electronic circuit diagrams for blind technicians and hobbyists. Tactile drawings often have to be hand made, but are increasingly produced with computer-driven Braille embossers. Touch tablet technology can be used to sense the position of the user's finger on a surface on which a raised tactile image is presented, and provide speech or audio feedback describing the image feature the user is touching. Active tactile displays up to 32 by 48 dots are available with an array of pins set to the "up" or "down" position under computer control to form a desired image. In some cases, graphs and oscilloscope waveforms can be coded auditorily so the pitch of a sound varies with the height of the graph or waveform as it is scanned from left to right.

Access to Appliances, Displays, and Daily Living Activities

On older appliances using control knobs, and on mechanical devices such as insulin syringes, tactile or large print labels can be applied to provide independent verification of settings. Modern appliances and instruments with low contrast LCD displays and complex menus make access by both blind and low vision users problematic. Therefore, assistive technology companies have produced talking versions of certain commonplace devices and appliances such as clocks, calculators, scales, and currency identifiers, blood-pressure monitors and sugar/insulin analyzers. However, there is still a need to impress on mainstream manufacturers the importance of considering the visually impaired when designing products.

Jobsite Adaptation Technology

Numerous special adaptations can be made to facilitate specific job-related tasks. Accessible meters, oscilloscopes, vacuum gauges, machinists' instruments, and all manner of vocational tools with auditory and tactile readouts have been developed. Jigs and fixtures can facilitate positioning, measurement, or the use of cutting tools, and improved lighting and magnification can help workers with low vision. In some cases, jobs can be restructured so task components that are critically dependent on vision are performed by another worker. Through these means, totally blind individuals have been employed in a wide array of jobs and professions.

Recreational Technology

Skiing (primarily cross-country), sailing, and kayaking are popular sports for blind people. Braille and auditory compasses are available, as are kayak and canoe paddles with tactile grips that help orient them correctly. Bowling is aided by a "bowling rail" that gives orientation to the bowling lane. Goal Ball is a game akin to soccer for blind or blindfolded players, in which the ball emits auditory signals to help localize its position. Playing cards and popular board games are available with large print, Braille, or tactile markings. Some blind people pursue electronics, both audio and amateur radio, as a hobby, using accessible audio and audio-tactile electrical meter readouts.

Audio descriptions for the visual action in movies and TV programs are available for certain PBS and cable television series and documentaries and on some home videos. Captioning and audio description of many new release movies are available in some theaters. For low-vision viewers, watching television and live performances can be aided with head-mounted monocular or binocular viewers, including electronic magnifying systems with head-mounted displays.

John Brabyn

See also Ageing and Vision; Audition: Disorders; Auditory Scene Analysis; Braille; Contrast Enhancement at Borders; Echolocation; Low Vision; Neuropsychology of Perception; Prostheses: Visual; Sensory Reconstruction and Substitution; Tactile Map Reading; Visual Disorders: Blindness

Further Readings

AFB Access World: An online magazine with reviews of current technology from the blind accessibility viewpoint. Retrieved from http://www.afb.org/aw/main.asp

Kendrick, D. (2000). *Business owners who are blind or visually impaired.* New York: AFB Press.

Seelman, K. D., Palmer, C.V., Ortmann, A., Mormer, E., Guthrie, O., Miele, J., & Brabyn, J. (2008, March–April). Quality-of-life technology for vision and hearing loss. *IEEE engineering in medicine and biology magazine, 27*(2), 40–55.

ATMOSPHERIC PHENOMENA

People hear some remarkable phenomena, such as thunder, and see some remarkable phenomena, such as sun rays, glories, and rainbows, in the atmosphere. These phenomena are studied mainly by physicists and involve the disciplines of *atmospheric acoustics* and *atmospheric* or *meteorological optics.*

Atmospheric phenomena are interesting to perceptionists because perception sometimes disagrees with reality (making these phenomena seem to be illusions) and because they give an opportunity to test the generality of perceptual principles discovered in the laboratory. This entry describes atmospheric acoustics and optics.

Atmospheric Acoustics

Atmospheric acoustics concern how sounds travel through the atmosphere. A notable phenomenon of sounds is the *Doppler effect.* As a sound source moves at constant speed, such as a vehicle with a siren, toward an observer (or as the observer moves toward a stationary sound), the frequency received by the observer when it first becomes audible is higher than the emitted frequency. As the sound source gets closer to the observer, the received frequency decreases until, closest to the observer, it equals the emitted frequency. As the sound source recedes from the observer, the received frequency becomes less than the emitted frequency. Many people, however, describe their experiences as an increase in pitch as the sound source approaches. This illusory initial increase in pitch might be from the increase in intensity of the approaching sound.

Another phenomenon is that distant sounds sometimes fluctuate in intensity markedly. For example, a single-engine plane flying some distance away sometimes sounds as if the pilot has cut the engine. This occurs because distant sounds can take more than one path, via inhomogeneities in the atmosphere, to the ears of an observer. Identical sounds traveling different paths can have their relative phases shifted, leading to cancellation when exactly out-of-phase waves arrive at the ear and to reinforcement when the waves are in-phase.

Atmospheric Optics

The atmosphere has its own optical properties. It can make its own light and it can act like a filter, a prism, and a mirror.

Atmosphere's Own Light

Apart from lightning, the atmosphere can make at least two other sorts of light, seen only at night. *Airglow* is weak light emitted by the upper atmosphere from chemical processes and from interactions between air molecules and cosmic rays. It is dim, about 0.3 ten-thousandths of a candle per square meter, visible only to rod vision as a slight colorless glow. It means, however, that even the darkest sky between the stars is not completely dark.

Aurorae are much stronger lights emitted by molecules of the upper atmosphere when bombarded by energetic particles from the sun. The particles follow the earth's magnetic field lines down the atmosphere, exciting oxygen molecules, giving red and green light, and nitrogen molecules, giving blue-violet light. Aurorae form in a ring between about 20 and 30° latitude from the magnetic poles.

Aurorae can be dim, visible only to rod vision, hence colorless. But other times, they can be bright enough to be seen in color by cone vision above other light sources such as the moon and streetlights. Aurorae can appear as diffuse glows or, startlingly, as vertical rays that look like curtains running approximately east-west. They can be essentially stationary or can form and reform over tens of seconds. The curtains can appear to be blowing about as if in a wind, parts overlapping with other parts. All this light and change occurs completely silently. Aurorae are some of the biggest things one can see in the atmosphere, being 80 kilometers (km) high and thousands of kilometers long. Occasionally the rays can descend directly over an observer, appearing as pattern of diverging rays from linear perspective.

Atmosphere as a Filter

To us, the atmosphere looks transparent. But this is because the wavelengths passed best by the atmosphere for identifying objects are just those to which we are most sensitive. The atmosphere largely protects us by filtering out harmful wavelengths emitted by the sun, such as microwaves and ultraviolet light. But even within the range of visible wavelengths, the sky can look blue or red, and the atmosphere can change the color and contrast of distant objects. These filtering properties arise from *scattering*. Moreover, clouds can be opaque, producing various *shadow* phenomena.

Scattering

About 17% of photons strike particles in the atmosphere, such as air molecules, and have their directions changed anywhere from close to 0°, *forward scattering,* to 180°, *back scattering.* Scattering ensures that the midday sky, away from the sun, is blue. This is because short-wavelength light that appears blue is scattered more than is longer-wavelength light that appears green, yellow, and red.

Scattered light, *skylight,* acts as a cue to the distance of objects. The further an object is from someone's eyes, the more skylight is scattered into his or her line of sight. This extra light reduces the contrast of markings on the object, reduces the contrast of the object with its background, and makes it bluish. If an object is far enough away, so much skylight will be superimposed on it will be indistinguishable from the background sky.

People use these properties of the atmosphere, *aerial perspective,* to judge distances (see color insert, Figure 11a). Aerial perspective does require some calibration: There are many stories of people from smoggy cities coming to some place of clean air and deciding to go for a half-hour walk to a hill that appears nearby, only to discover it is a mountain scores of kilometers away.

Shadows

Clouds consist of numerous particles of water or ice. If they are thick enough, essentially all of the light that falls on them is reflected, casting a shadow on the opposite side from the sun. For a viewer looking at a cloud with the sun behind it, the cloud can look dark. Near the edge, where the cloud is thinner, a viewer can sometimes see a brilliant silver or golden lining from forward-scattered light. Clouds look brightest, whitest, when the sun is behind the viewer. But sometimes in this situation one sees a dark cloud against a more distant white cloud. This is partly from geometry and partly from perception. The dark part is the underside of a near cloud in shadow. The white part is the top of distant clouds in sun (Figure 1). Aerial perspective also superimposes skylight on distant clouds, lightening them. Perception adds *simultaneous brightness contrast,* in which a gray area on the light background of another cloud looks darker than an identical gray area on the dark background of the blue sky.

Cloud shadows also create dark parts of the atmosphere. These are easiest to see when the sun is setting in the west behind clouds and are revealed as a fan-shaped pattern of dark and bright rays that appear to diverge from the sun—*crepuscular rays.* But these rays do not connect with, or diverge from, the sun; they exist only in the atmosphere where there are particles to scatter light into a viewer's eyes. And because the sun's light is parallel, the rays are parallel too. The rays appear to diverge from linear perspective because they are approaching the viewer.

Occasionally crepuscular rays can be seen continuing over the bowl of the sky, following great circles, and then converging to the antisolar point in the east: exactly opposite to the position of the sun in the western sky. These *antisolar crepuscular rays* or *anticrepuscular rays* (see color insert, Figure 11b) are also a simple product of geometry and linear perspective. The rays appear to converge because they are really parallel and receding from the viewer.

Atmosphere as a Prism or as an Interference Filter

A prism *refracts* light, bends it. Because light is bent by an angle related to the wavelength, with short wavelengths being bent more than long wavelengths, prisms also break sunlight into its component wavelengths, called *dispersion.* Refraction occurs in air because air's density varies (giving rise to mirages), and it occurs in ice crystals (giving various halos) and in water drops (giving rainbows).

Some of the beautiful spectral colors in the atmosphere arise via interference between light waves from differences in their phase. These are commonly seen in thin films, such as in soap bubbles, or in interference filters, such as found fortuitously on a compact disk. An interference filter usually consists of a grating of fine slits through which light can pass. The phase differences arise from light from one slit traveling a slightly different distance to the eye than light from another slit. When a particular wavelength is reinforced, it superimposes its associated color onto whatever light is entering the eyes. Interference phenomena occur in the atmosphere from droplets of water or from ice needles, each element acting like a slit in an interference filter (giving *iridescent clouds* and *coronae*).

Atmosphere as a Mirror

The atmosphere can behave like a mirror from backscattering of small particles at angles approximately equal to 180°, giving rise to *glories* and the *Brocken spectre,* and from ordinary reflection from the surfaces of water droplets or ice crystals or from total internal reflection within droplets or crystals, giving rise to *pillars* and to *rainbows.*

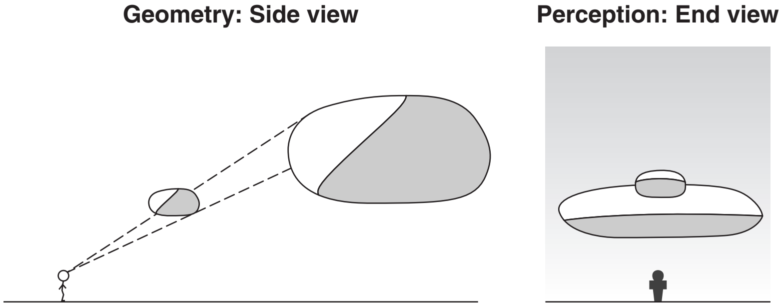


Figure 1 Dark Versus White Clouds

Notes: (Left) Side view of the geometrical arrangement of clouds producing a dark, near cloud visible against a white, distant cloud. The dashed lines show a viewer's lines of sight. (Right) End view of how a dark cloud appears against a white cloud. The shadowed parts of the distant cloud can also be lighter than those of the near cloud because of aerial perspective.

Glories and the Brocken Spectre

Small drops of water, such as in fog or mist, have pronounced backscattering, which is diffuse reflection. When the drops are uniform, they can produce colored rings from interference. These phenomena require strong sunlight to illuminate a viewer and then to pass into the drops; these conditions were originally the preserve of mountain climbers, often seen from the Brocken, an easily accessible peak in Germany's Harz Mountains. Now they are usually seen in clouds by passengers in aircraft. Light blocked by the viewer's head or aircraft creates the *specter,* a shadowed region of fog or cloud; light just missing the obstruction is backscattered to create a bright halo. Surrounding this can be at least one set of colored rings, the *glory.* Often, a viewer's shadow appears magnified, presumably from overestimation of its distance in fog (see color insert, Figure 11b).

Pillars and Rainbows

True reflection can occur when there is an abrupt, large change in the refractive index from one part of the atmosphere to another. This commonly happens on the flat surfaces of ice crystals and in drops of water. When ice crystals float in the air with their axes mainly horizontal, light can be reflected from their surfaces into a viewer's eye, seen as a vertical ray or *pillar* passing through the light source. These reflections have the same color as the light. Drops of water are approximately spherical, so reflection from the surface is weak. But rays of light can enter a drop at one point, then refract to meet an opposite surface at an angle giving *total internal reflection.* The rays also undergo dispersion within the drop, yielding the most spectacular examples of refraction colors: the rainbow (see color insert, Figure 11c).

All these phenomena share some perceptual properties. Each person sees his or her own specter, glories, halo, coronae, and rainbows because only light from the correct angle will pass into a viewer's eyes to create these phenomena. This is most startling in the case of specters and glories because these phenomena can be only 10 to 100 meters from a viewer. A man and a woman, side by side at the top of a mountain, will each see his or her own specter, the woman's being invisible to the man, and the man's being invisible to the woman. It is the same, but less noticeable, with the other phenomena. For example, with rainbows, one person's red region of a bow could be a neighbor's blue region.

All the phenomena have constant visual angles, the angle at the apex of a triangle at the eye and its base at the object. Unlike objects, whose visual angles decline with distance from the eye, these phenomena appear only when rays at the correct angle enter the eye. For example, the visual angle of the width of a rainbow is about 2°; it has this angle whether it is made in the spray from a nearby hose or in distant rain. Consistent with geometry of objects, the rainbow in the hose looks small whereas that in the rain looks enormous. The perceived size is determined by perceived distance.

The perceived distance of these phenomena is mainly determined by the physical distance from a viewer's eyes of the parts of the atmosphere producing them. But it will be influenced by other depth information. For example, a rainbow's end with a visual direction similar to some conspicuous geographical feature (such as a headland jutting into the sea) will tend to have the same depth as that feature. But of course, as soon as one goes to the foot of the rainbow, to find the legendary pot of gold there, it disappears, leaving the searcher only wet.

Robert P. O'Shea

See also Air Quality; Audition: Pitch Perception; Audition: Temporal Factors; Auditory Illusions; Color Constancy; Color Perception; Constancy; Depth Perception in Pictures/Film; Lightning and Thunder; Mirages; Nonveridical Perception; Visual Illusions

Further Readings

Cowley, L. (n.d.). *Atmospheric optics.* Retrieved March 20, 2008, from http://www.atoptics.co.uk

Lynch, D. K., & Livingston, W. (2001). *Color and light in nature* (2nd ed.). Cambridge, UK: Cambridge University Press.

Minnaert, M. G. J. (1954). *The nature of light and colour in the open air* (H. M. Kremer-Priest, Trans.). New York: Dover.

ATTENTION: COGNITIVE INFLUENCES

In 1880, William James famously defined *attention* as the mind taking possession of one out of several simultaneously possible objects or train of thoughts. The modern study of attention continues to work within this broad definition. It is generally accepted that an individual is only aware of a small fraction of the information provided to the brain by the sensory systems. *Attention* is the name given to the process that governs which material enters awareness and which does not. This entry focuses on how cognitive factors—goals and expectations of an observer—influence visual attention.

Whether a given stimulus is attended depends both on its inherent salience and the state of the observer. The interplay between these two factors is exemplified by the following two situations: (1) Imagine searching for a friend in a crowd. As one searches through the crowd, attention may be captured by elements that are inherently salient, for instance, a person in a red coat among people in black coats, or by dynamic cues such as a person running. However, where one looks in the crowd also depends on one's knowledge. So, if the friend is known to be wearing a black coat and a blue hat, attention may be less likely to be captured by an otherwise salient red coat, but perhaps be misdirected to a blue coat as one searches for anything blue in the crowd. (2) Imagine walking by a golf course and worrying about being hit by a golf ball. Although it is likely that a quickly moving white thing in one's visual field will be noticed under most circumstances, the act of walking by a golf course and worrying about being hit by a golf ball may lower the threshold for detecting quickly moving white things. Directing one's attention to blue things because blue is currently relevant and increasing sensitivity to detecting moving white things are instances of *cognitive influences on attention.* Traditional accounts of attention have placed little emphasis on such influences in comparison with factors that were thought to automatically capture attention (e.g., a red thing among black things). However, it can be argued that the purpose of visual attention is the selection of information most relevant to a present goal. Hence, the effects of goals, expectations, recent history, and even emotions on visual attention have become an area of active research.

Studying Attention: Behavioral Methods

A commonly used paradigm for studying visual attention is the eponymously named Posner cueing task. The basic version of the task requires subjects to press a button anytime they detect a small circle (the target) while looking at the center of a screen without moving their eyes. Before the appearance of the target, a light (the cue) flashes on the left or the right side of the screen. The location of the cue either coincides with the location of the subsequently appearing target (valid trials) or does not (invalid trials). The basic finding is that reaction times are shorter on valid trials than on uncued trials, and are slowest on invalid trials. The interpretation is that attention is automatically "deployed" to the cued region. Targets that appear in the attended region are processed faster than are targets appearing in an unattended region (which require an attentional shift from the previously cued region).

In addition to flashing lights, the cues can be symbolic, such as right or left arrows presented in the center of the screen. Classic studies from the late 1970s showed that flashing lights (also called exogenous or peripheral cues) elicit attentional shifts even when they do not predict the location of the target. In contrast, arrow cues (also called endogenous or central cues) only produce shifts of attention when they are predictive (e.g., the arrow predicts the position of the target 80% of the time). These results have been interpreted to mean that, unlike flashing lights that automatically capture attention, central cues need to be cognitively interpreted and will shift attention only if subjects have a reason to process them. These classic findings led to a dichotomy between automatic (stimulus-driven, bottom-up, exogenous) attentional processes and controlled (cognitive, top-down, endogenous) processes.

Controversies

Recent studies have argued against this dichotomy in favor of a view in which (a) learned associations determine whether nonpredictive endogenous cues elicit attentional shifts, and (b) highly salient cues can fail to capture attention if they conflict with the viewer's goals. For example, studies from Alan Kingstone's laboratory have shown that pictures of eyes elicit attentional shifts in the direction of their gaze even when the direction does not predict the target. Similarly, nonpredictive arrows and even printed words such as *up* and *down* elicit attentional shifts. Conversely, whether a traditional exogenous cue such as a unique color (e.g., a patch of red among greens) captures attention appears to depend on how the viewer is processing the scene. Consider performance on a task developed by Jan Theeuwes to study the degree to which various visual properties automatically capture attention. In this task, participants are presented with shapes arranged on an imaginary circle around a central fixation point. The goal is to report whether a line that appears in a target shape is, for example, vertically oriented. In a basic version of this task, the target is defined by its unique shape (e.g., it is the only diamond among circles). On distractor trials, the display appears with one of the nontarget shapes in a different color from the rest. Because color is irrelevant to the task, greater reaction times on distractor trials indicate that the unique color automatically captured attention. Based on such findings, Theeuwes and colleagues argued that uniquely colored or shaped objects (singletons) automatically capture attention. Howard Egeth and colleagues challenged this conclusion by showing that whether unique objects capture attention depends on the processing mode of the viewer. If the viewer is in a "singleton detection mode," tuned to detect unique objects, attention is broadly focused and is indeed captured by task-irrelevant singletons. However, if the viewer is specifically looking for a certain feature such as a diamond shape (*feature detection mode*) then salient but task-irrelevant distractors do not capture attention. Nevertheless, there do exist properties that capture attention regardless of task relevance or processing mode. One such property is *sudden onset*. A suddenly appearing object generally captures attention. However, when the object is task-irrelevant, attention is disengaged quite quickly (typically in less than 100 milliseconds [ms]).

Although there is now wide agreement that the viewers' goals can affect which objects or features are attended, the locus of these effects remains highly controversial. Does having a goal like "look for the red things" change the priority of redness but not affect visual processing? Or, does the goal actually change how red things are represented throughout the visual system? Traditional accounts have denied the latter claim. For example, Zenon Pylyshyn has argued for the existence of an early vision system—a modular system that is encapsulated from information outside vision such as the observer's knowledge and goals, and is thus "cognitive impenetrable." Support for the claim that attention changes basic visual processes has come from behavioral studies showing that attended objects are actually perceived as more salient (e.g., brighter), from electrophysiological studies on nonhuman animals and neuroimaging studies on humans.

Effects of Goals on Visual Processing and Attention: Evidence From Electrophysiology and Neuroimaging

Electrophysiological studies on behaving animals have allowed researchers to isolate higher-level influences on attention from the processing that reflects the physical properties of the stimuli. For instance, because neurons in the primary visual cortex (V1) fire most to bars with a certain orientation, a researcher can compare the firing of the neuron to a vertical bar when it is task-relevant versus irrelevant. Any difference in neuron's firing rate to a particularly tuned bar between the first and second task reflects the demands of the task because the physical stimulus is identical in both cases. Such studies generally show that V1 neurons fire more vigorously when their preferred orientation is behaviorally relevant.

An immediate implication of such findings is that responses of neurons even in V1—the first part of the cortex to receive visual input—reflect not simply the physical characteristics of a stimulus, but also the cognitive goals of the observer. More recent findings have shown that primary sensory neurons have two types of receptive fields (RFs). The first is the so-called classical RF and corresponds roughly to what is observed in anesthetized animals and is the initial response of a neuron in an awake animal. The classical RF of a V1 neuron is a line segment of a certain orientation projected into a specific part of the visual field. Within a short period (often less than 50 ms, and sometimes as short as 2 ms), the classical RF is modulated by higher-level information including the goals of the observer, the visual context, and the organization of the scene—producing the nonclassical RF. Although the classical RF of a V1 neuron includes only positional and orientation information, the nonclassical RF includes information such as whether the bar is part of a figure, the background, or an object boundary and whether the figure that the segment is a part of is behaviorally relevant.

In similar studies measuring firing rates of Area V4 neurons (sensitive to color properties of a stimulus), attentional capture of task-irrelevant color singletons is reflected in high firing rates that peak at approximately 120 ms after stimulus onset. When the task requires the monkey to ignore the color singleton, one can observe neural responses to a task-relevant color continue to remain at a high level, whereas responses to task-irrelevant color singletons become down-modulated after approximately 75 ms.

Is early visual neural activity immune to cognitive influences? In electrophysiological studies, it has been found that when a task is performed repeatedly, such as attending to a vertical bar for numerous consecutive trials, the classical RF may disappear entirely—the neurons' response being immediately modulated by the current task. Recent neuroimaging work in humans confirms the conclusion that activity in anatomically early visual areas is permeable to cognitive influences. For instance, when human observers are trained to associate cues with either color or attention, presenting the cue alone modulates activity in visual areas (fusiform gyrus for color; lingual gyrus for location) and, crucially, the amount of modulation strongly predicts performance on the upcoming target-detection trials. This suggests that early visual processing can be tuned by goals and expectations.

Recall that attention is thought to be closely linked to awareness (one notable exception is the phenomenon of blindsight). A claim that a red circle among black circles automatically captures attention generally means that one becomes aware of the red circle even if one's goal is to avoid it. Would showing that some early neural activity evoked by the red circle is impermeable to top-down influences be evidence that automatic awareness of the red circle is directly subserved by that early neural activity? Recent work suggests that such a conclusion is unwarranted. Rather, it appears there may be no awareness without top-down modulation of early visual representations. Much of this evidence has come from studies relying on event-related potentials (ERPs) transcranial magnetic stimulation (TMS). Both methods rely on the earlier time-course of bottom-up versus the later time-course of top-down processes to map their respective contributions. In one study, subjects detected visual figures that were briefly presented and then concealed with a pattern mask. By correlating subjects' performance with electrical potential measured by electrodes on the scalp, J. Fahrenfort and colleagues showed that bottom-up activity, which peaked at about 120 ms after stimulus onset, was not correlated with conscious perception, but top-down (recurrent) activity peaking later (160 ms) was. Thus, it appears that top-down modulation in anatomically "early" visual areas (e.g., V1) by higher-level regions (e.g., prefrontal cortex) is necessary for visual awareness. Further evidence for the causal role of recurrent processing comes from TMS studies in which a high-intensity magnetic field is briefly applied to a selected region of a subject's scalp as he or she performs a task. This pulse creates a temporary disruption in neural processing. Several recent studies have shown that disrupting feedback activity in early visual areas disrupts awareness.

Research Conclusions

The study of attention has classically focused on the physical characteristics that determine whether stimuli are attended. Recent studies have shifted the focus to cognitive factors such as expectations and goals of the viewer. These studies show that neural activity causally linked to awareness is deeply permeated by cognitive factors. Although highly controversial, one conclusion is that it may be impossible to fully study visual attention by separating observers from their goals and environments.

Gary Lupyan

See also Attention: Physiological; Attention and Consciousness; Neural Representation/Coding; Nonveridical Perception; Object Perception

Further Readings

Fahrenfort, J., Scholte, H., & Lamme, V. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of Cognitive Neuroscience, 19*(9), 1488–1497.

Foxe, J., & Simpson, G. (2002). Flow of activation from V1 to frontal cortex in humans—A framework for defining "early" visual processing. *Experimental Brain Research, 142*(1), 139–150.

Giesbrecht, B., Weissman, D. H., Woldorff, M. G., & Mangun, G. R. (2006). Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. *Brain Research, 1080*(1), 63–72. DOI: 10.1016/j.brainres.2005.09.068.

Kanwisher, N., & Wojciulik, E. (2000). Visual attention: Insights from brain imaging. *Nature reviews. Neuroscience, 1*(2), 91–100. DOI: 10.1038/35039043.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., & Eastwood, J. D. (2003). Attention, researchers! It is time to take a look at the real world. *Current Directions in Psychological Science, 12*(5), 176–180. DOI: 10.1111/1467–8721.01255.

Lamme, V., & Roelfsema, P. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences, 23*(11), 571–579.

Lamy, D., & Egeth, H. E. (2003). Attentional capture in singleton-detection and feature-search modes. *Journal of Experimental Psychology. Human Perception and Performance, 29*(5), 1003–1020. DOI: 10.1037/0096–1523.29.5.1003.

Posner, M., Snyder, C., & Davidson, B. (1980). Attention and the detection of signals. *Journal of Experimental Psychology–General, 109*(2), 160–174.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22*(3), 341–364.

Theeuwes, J., & Van der Burg, E. (2007). The role of spatial and nonspatial information in visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 33*(6), 1335–1351. DOI: 10.1037/0096–1523.33.6.1335.

ATTENTION: COVERT

Each time we open our eyes, we are confronted with an overwhelming amount of information. Despite this, we have the clear impression of understanding what we see. This requires selecting relevant information out of the irrelevant noise. Attention is what turns looking into seeing, allowing us to select a certain location or aspect of the visual scene and to prioritize its processing. Such selection is necessary because the limits on our capacity to absorb visual information are severe.

They may be imposed by the fact that there is a fixed amount of overall energy consumption available to the brain, and by the high-energy cost of the neuronal activity involved in cortical computation. Attention is crucial in optimizing the use of the system's limited resources, by enhancing the representation of the relevant locations or features while diminishing the representation of the less relevant locations or aspects of our visual environment.

The processing of sensory input is facilitated by knowledge and assumptions about the world, by the behavioral state of the organism, and by the (sudden) appearance of possibly relevant information in the environment. For example, spotting a friend in a crowd is much easier if you are cued to two types of information: where to look and what to look for. Indeed, numerous studies have shown that directing attention to a spatial location or to distinguishing features of a target can enhance its discriminability and the neural response it evokes. Understanding the nature of attention and its neural basis is one of the central goals of cognition, perception, and cognitive neuroscience, as described in this entry.

Spatial Covert Attention

Attention can be allocated by moving one's eyes toward a location, overt attention, or by attending to an area in the periphery without actually directing one's gaze toward it. This peripheral deployment of attention, known as covert attention, aids us in monitoring the environment and can inform subsequent eye movements. Cognitive, psychophysical, electrophysiological, and neuroimaging studies provide evidence for the existence of overt and covert attention in both humans (including infants) and nonhuman primates. Many of these studies have likened attention to increasing visual salience. Whereas covert attention can be deployed to more than one location simultaneously ("in parallel"), eye movements are necessarily sequential ("serial"); they can only be at one location at a given time. Many studies have investigated the interaction of overt and covert attention and the order in which they are deployed. The consensus is that covert attention precedes eye movements and their effects on perception, which in many cases are similar but in others they are not.

Hermann von Helmholtz is considered to be the first scientist to provide an experimental demonstration of covert attention (circa 1860). He experimented with a wooden box whose interior was completely dark. Looking into the box through two pinholes, he reported that he could concentrate on any part of the visual field so that when a spark came, he could focus attention independently of the position and accommodation of his eyes, and get an impression of objects in only the particular attended region. The similarities and differences between eye movements and deployment of spatial attention have been a focus for research ever since.

Many authors state that humans deploy covert attention routinely in many everyday situations, such as searching for objects, driving, crossing the street, playing sports, and dancing. However, other authors think that covert attention is deployed mainly in social situations—for example, when deception about intentions is desired, in competitive situations (such as sports activities), or when moving the eyes would provide a cue to intentions that the individual wishes to conceal.

Endogenous and Exogenous Covert Attention Systems

A growing body of behavioral evidence demonstrates that there are two covert attention systems that deal with facilitation and selection of information: *endogenous* and *exogenous*. The former is a voluntary system that corresponds to our ability to willfully monitor information at a given location; the latter is an involuntary system that corresponds to an automatic orienting response to a location where sudden stimulation has occurred. Endogenous attention is also known as *sustained* attention and exogenous attention is also known as *transient* attention. These terms refer to temporal nature of each type of attention: Whereas observers seem to be able to sustain the voluntary deployment of attention to a given location for as long as needed to perform the task, the involuntary deployment of attention is transient, meaning it rises and decays quickly. The different temporal characteristics and degrees of automaticity of these systems suggest that they may have evolved for different purposes and at different times—the transient, exogenous system may be phylogenetically older.

Behavioral Studies

To investigate covert attention, it is necessary to make sure that observers' fixation is maintained, and to keep both the task and stimuli constant across conditions while manipulating attention. Experimentally, the endogenous and exogenous systems can be differentially engaged by using distinct cues. Michael Posner devised a paradigm that has been widely used to study the endogenous and exogenous orienting of attention. It allows the comparison of performance in conditions where attention is deliberately directed to either a given location (attended condition), away from that location (unattended condition), or distributed across the display (neutral or control condition).

In the Posner cueing paradigm, observers have to respond as quickly as possible to a peripheral target, which is preceded by a central or peripheral cue. In the endogenous condition, a central cue—typically an arrow—points to the most likely location of the subsequent target. In the exogenous condition, a brief peripheral cue is typically presented at one of the target locations but does not predict the subsequent target location. In both endogenous and exogenous conditions, performance in detecting or discriminating a target is typically better (faster, more accurate or both) in trials in which the target appears at the cued location than at uncued locations. Central or symbolic cues take about 300 milliseconds (ms) to direct endogenous attention in a goal-driven or conceptually driven fashion. Central cues are small lines presented at fixation pointing to different locations of the visual field (e.g., upper left quadrant, lower right quadrant); symbolic cues include different numbers or colors that indicate different locations where the observer is to attend (e.g., #1 or a red circle indicates upper left quadrant, #3 or a blue circle indicates lower right quadrant). Peripheral cues, conversely, grab exogenous attention in a stimulus-driven, automatic manner within about 100 ms. Whereas the shifts of attention elicited by central cues appear to be under conscious control, it seems that it is extremely hard for observers to ignore peripheral cues. This involuntary transient shift occurs even when the cues are uninformative or may impair performance.

Neurophysiological Studies

Single-Unit Recording Studies

The development of techniques to record the electrical activity of single neurons in awake-behaving animals (e.g., monkeys) has enabled researchers to probe the biological foundations of endogenous (sustained) attention while monkeys perform attention-demanding tasks. Such studies have provided detailed, quantitative descriptions of how endogenous attention alters the responses of neurons in the extrastriate visual cortex, yielding attentional facilitation and selection. Attentional facilitation results when spatial attention enhances the responses evoked by a single stimulus appearing alone in a neuron's receptive field, so that neurons respond to an attended stimulus much as they would were its luminance increased. Given that stimuli rarely appear in isolation, attentional selection of behaviorally relevant targets from among distracters arguably serves a more ecologically relevant purpose. When multiple stimuli appear within a neuron's receptive field, the firing rate is characteristically determined primarily by the task-relevant stimulus. Numerous studies have compared the response when attention is directed either with one of the two stimuli in the receptive field or outside the receptive field while fixation is maintained. Attending to the preferred stimulus (for which the neuron is tuned) increases the neuron's response evoked by the pair of stimuli whereas attending to the non-preferred stimulus (for which the neuron is not tuned) decreases such response.

Neuroimaging Studies

Neuroimaging has yielded information on the integrated brain activity underlying perception and cognition in humans. Studies documenting the neural correlates of covert attention have used several techniques, among them event-related potentials (ERPs) and functional magnetic resonance imaging (fMRI).

ERPs are electrophysiological responses that arise during sensory, cognitive, and motor processing, which provide precise information about the time course of information processing. ERP recordings can help reveal the timing and organization of stimulus selection processes in the brain's attentional network. ERP studies provide support for a mechanism of early sensory facilitation, at the level of the extrastriate visual cortex, during the spatial cueing of attention.

fMRI measures hemodynamic processes not invasively in the human brain by providing temporally integrated maps of regional cerebral blood flow across the whole brain. It is based on the increase in blood flow to the local vasculature that accompanies neural activity in the brain. There is wide agreement that attention increases fMRI responses in visual cortical areas in a retinotopically specific manner, corresponding to attended spatial locations, both for endogenous and exogenous attention. There is no consensus, however, about whether common neurophysiological substrates underlie endogenous and exogenous attention. Whereas some fMRI studies have found no difference in the brain networks mediating these systems, others have reported differences. For example, endogenous attention is cortical in nature, but exogenous attention also activates subcortical structures. Moreover, partially segregated networks mediate the preparatory control signals of both systems. For example, endogenous attention is mediated by a feedback mechanism involving feedback from frontal and parietal areas, whereas these regions are not necessarily involved in exogenous attention.

Feature-Based Attention

Behavioral Studies

Most studies of overt or covert attention (with or without eye movements, respectively) have examined the effects of spatial attention at different locations in the visual field. Both types of attention, however, can also be allocated to specific visual features, such as orientation or direction of motion or color. Feature-based attention is the ability to enhance the representation of image components throughout the visual field that are related to a particular feature. This type of attention is considered to play a central role when human or nonhuman primates search for a stimulus containing that feature. This ability to detect a target or relevant item among distracters is the basis of a widely used paradigm in visual attention research: visual search.

Many psychophysical studies have shown that feature-based attention improves detection and discrimination performance across the visual field. A visual stimulus always occupies a certain spatial location, so it is important to control spatial selection. Thus, studies of feature-based attention generally use compound stimuli that contain multiple features superimposed over the same spatial location, and observers are required to attend to one of those features. For instance, attending to one motion direction in a compound motion stimulus produces a motion aftereffect consistent with the attended direction, and attending to one orientation in a compound orientation stimulus produces an orientation aftereffect consistent with the attended orientation.

Neurophysiological Studies

Because feature-based attention is independent of location, it is well suited to selectively modify the neural representations of stimulus features within visual scenes that match the currently attended feature. For example, feature-based attention can enhance the neural responses to a given attended feature; e.g., vertical orientation regardless of the locations where vertical orientations appear in the scene.

Single-Unit Recording

Some studies using single-unit recordings to investigate feature-based attention have shown that the responses of individual neurons that are selective for an attended feature (for example, vertical orientation, upward motion direction, or red color) can be selectively enhanced when the monkey attends to that particular feature (for example, upward motion while seeing a stimulus containing both upward and downward motion). Other studies have shown that attending to different feature dimensions (e.g., color or orientation) modulates activity in cortical areas specialized for processing those dimensions.

Neuroimaging

fMRI studies requiring participants to attend to specific features of stimuli have revealed that the neural activity evoked by the attended feature is enhanced both within the retinotopic cortical regions representing the spatially attended location and at other unattended regions that were stimulated by the same feature. Enhanced activation has even been observed at spatial locations where no stimuli were present. These studies suggest that feature-based attention is activated across the visual field. This representation enables the visual system to concentrate its limited processing resources on the most relevant sensory inputs regardless of where in the visual field they are located.

Marisa Carrasco

See also Attention: Effect on Perception; Attention: Object-Based; Attention: Physiological; Attention: Selective; Attention: Spatial; Attention: Theories of; Eye Movements: Behavioral; Visual Search

Further Readings

Carrasco, M. (2006). Covert attention increases contrast sensitivity: Psychophysical, neurophysiological, and neuroimaging studies. In S. Martinez-Conde, S. L. Macknik, L. M. Martinez, J. M. Alonso, & P. U. Tse (Eds.), *Visual perception. Part I. Fundamentals of vision: Low and mid-level processes in perception* (pp. 33–70). Amsterdam: Elsevier.

Corbetta, M., & Shuiman, G. L. (2003). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3,* 201–215.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29,* 1631–1647.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Reynolds J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annual Review of Neuroscience, 27,* 611–647.

ATTENTION: CROSS-MODAL

Attention refers to those processes that allow for the *selective* processing of incoming sensory stimuli, typically those that are most relevant to one's current goals or to the task at hand, or, alternatively, those that have the greatest intrinsic salience or biological significance. Attended stimuli tend to be processed both more thoroughly and more rapidly than are other potentially distracting ("unattended") stimuli. Although research on attention has traditionally considered selection among the competing sensory inputs within just a single modality at a time (most often vision), the past two decades have seen a burgeoning of interest in the existence and nature of any cross-modal constraints on our ability to selectively attend to a particular object, location, or source of information.

The term *cross-modal* is typically used in situations in which the orienting of a person's spatial attention in one sensory modality (such as vision) results in a concomitant shift of attention in one or more of his or her other sensory modalities (such as audition or touch) to the same location (or object) at the same time. The central question for researchers interested in cross-modal attention concerns how the brain's attentional resources are coordinated or linked between the senses to select just that subset of information that is relevant to a person's current goals from among the abundance of multisensory information impinging on the various sensory receptors at any one time. This entry describes coordination of attentional resources, attentional selection research, mapping of cross-modal links, and future research directions

Coordination of Attentional Resources

One intuitive possibility is that there might be independent modality-specific attentional resources. So, for example, some researchers have argued that there may actually be separate visual, auditory, and tactile attentional systems in the human brain. According to this account (which posits that cross-modal links in spatial attention do not exist), people should, for example, be able to direct their visual attention to one location while directing their auditory or tactile attention to a different location (because the attentional systems are independent). However, other researchers have argued that there is only a single supramodal attentional system, such that people can only attend to a single location at any given time (that is, they cannot split their attention between different locations simultaneously). According to the supramodal account, all stimuli, no matter what their modality, that are presented from a location that is attended should receive preferential processing over stimuli that are presented elsewhere.

A third possibility (that has emerged more recently) is that there might be some intermediate form of organization instead. So, for example, according to Charles Spence and Jon Driver's hybrid separate-but-linked hypothesis, there may be separate visual, auditory, and tactile attentional systems at the earliest levels of human information processing. However, these attentional systems are subsequently linked, such that people's attention typically tends to be (but importantly does not always have to be) focused on the same location in space in the different modalities.

Attentional Selection

Researchers studying spatial attention distinguish between endogenous and exogenous attentional selection. *Endogenous* attention is typically involved in the voluntary orienting of attention to a particular event or spatial location, such as when you choose to attend to a particular person at a noisy cocktail party, or when you concentrate on the feel of the object that you are playing with in your right hand. By contrast, *exogenous* (or involuntary) orienting occurs when attention is reflexively shifted to the location of a sudden and unexpected peripheral event, such as when a person calls your name at the cocktail party, or when a fly suddenly lands on your arm. Orthogonal to this distinction between endogenous and exogenous attention is the distinction between overt and covert attentional orienting: *Overt* orienting refers to shifts of receptors (as in eye, head, or hand movements), whereas *covert* orienting (which is of most interest to cognitive psychologists studying selective attention) refers to internal shifts of attention (e.g., as when we look at someone out of the corner of our eye).

Exogenous Attention

Most studies of cross-modal spatial attention have adapted the spatial cueing paradigm first popularized by Michael Posner back in the 1970s. In a typical study of exogenous attention, a cue stimulus is presented in one sensory modality (e.g., audition, vision, or touch) shortly before a target appears on either the same or opposite side of central fixation. The target stimulus may be presented in either the same or different sensory modality as the cue. Importantly, however, the target is just as likely to be presented on the same, as on the opposite, side as the cue (that is, the cue is spatially nonpredictive with regard to the likely target location). Participants are often instructed to try to ignore the cue as much as possible. However, the results of many studies conducted during the past 20 years have shown that participants typically cannot ignore such cues and respond more rapidly (or accurately) to targets presented at the cued, as opposed to the uncued, location, at least when the target is presented within a few hundred milliseconds of the cue. Depending on the particular task, participants sometimes start to respond more slowly to targets at the cued (compared with the uncued) location as the interval between the onset of the cue and target lengthens, a phenomenon known as *inhibition of return.*

The presentation of either an auditory or a tactile spatially nonpredictive cue has been shown to result in a rapid exogenous shift of spatial attention to the cued location, and this shift of attention facilitates the subsequent processing of auditory, tactile, and visual targets at that location. Meanwhile, visual cues have been shown to elicit a shift of visual, tactile, and, on occasion, auditory attention toward the cued location. Recent studies of exogenous cross-modal spatial attention have shown that perceptual sensitivity is enhanced at the cued location, and what is more, people tend to become aware of stimuli presented at the cued location sooner than when the same stimuli are presented elsewhere, a phenomenon known as *prior entry.* At longer cue-target intervals (typically greater than 200–300 milliseconds), inhibition of return has now been observed in speeded detection tasks between all combinations of auditory, visual, and tactile cue and target stimuli.

Endogenous Attention

Turning to the nature of the cross-modal links that constrain the deployment of endogenous spatial attention, researchers have shown that if people deliberately direct their attention to a particular location in one sensory modality, their endogenous attention in the other modalities will tend to follow to the same location, albeit at a somewhat reduced level (i.e., the attentional benefits are smaller). So, for example, if participants are instructed to attend to their left hand because a tactile target is likely there, visual targets will also be responded to preferentially by the left hand. Evidence in support of the separate-but-linked hypothesis of cross-modal links in endogenous spatial attention comes from the results of studies showing that although they find it difficult, people can nevertheless still direct their attention in different directions in different modalities at the same time. So, for example, under the appropriate experimental conditions it has been shown that people can preferentially process visual stimuli presented by their left hand, while showing an attentional bias toward the processing of tactile stimuli presented to their right hand. Results such as these are inconsistent with both the modality-specific and supramodal accounts of cross-modal spatial attention.

Mapping of Cross-Modal Links

Having demonstrated the existence of cross-modal links in both exogenous and endogenous spatial attention (and in inhibition of return) between all possible combinations of auditory, visual, and tactile stimuli, one of the most important issues currently in cross-modal research is how (and even whether) the brain updates the mapping (or correspondence) between the senses when people change their posture. Each of our senses initially codes information according to a different frame of reference: So, for example, visual stimuli are initially coded retinotopically, auditory stimuli tonotopically, and tactile stimuli somatotopically. The question therefore arises how all of this information is coordinated into a common frame of reference for the control of attention (and subsequently action). To investigate whether cross-modal links in spatial attention are updated following changes in posture, researchers have conducted experiments in which the participants have had to cross their hands over the midline (so that their left hand is in the right side of space and their right hand in the left), or to deviate their gaze (to either the left or right) while keeping their head fixed straight ahead. The results of several such studies now show that cross-modal links in spatial attention are updated following posture change. So, for example, researchers have shown that although people find it easier to attend to tactile stimuli presented to their left hand and to visual stimuli on the *left* side of space when their hands are uncrossed (so that their left hand is on the left side of their body), they find it easier to concentrate on their left hand and *right* visual stimuli when their hands are crossed over the midline (such that their left hand now lies in the right hemispace). Results such as these have led researchers to conclude that the "space" in which attention is directed is itself a multisensory construction.

Many researchers are currently investigating how (and where) such cross-modal links in spatial attention are mediated in the human brain. One popular suggestion has been that multisensory spatial maps, such as those found in the superior colliculus (where spatially aligned maps of visual, auditory, and tactile space are superimposed), might mediate at least some of the behavioral spatial cueing effects that have been observed in the laboratory. However, there is currently much debate about the extent to which exogenous cross-modal spatial attention effects (typically observed by cognitive psychologists in awake human participants) and multisensory integration effects (typically observed at the single cell level in anaesthetized animals by neurophysiologists) actually represent the same underlying phenomenon. There is, then, still much work to be done in bridging the gap between the different methods and the different levels of analysis at which scientists study cross-modal spatial attention. One popular idea that has emerged recently is that multisensory influences on unimodal brain areas might arise because of feedback or back-projection influences on them, from multisensory convergence-zones or attentional control structures.

Future Directions

Our growing understanding of the nature of the cross-modal links that constrain spatial attention will likely increasingly help provide guidelines to aid the effective design of multimodal (or multisensory) user interfaces. For example, research in the field of applied psychology has shown that people find it particularly difficult to hold a conversation on a mobile phone while driving a car. One of the major problems in this multisensory dual-task situation may be that people find it difficult to attend visually out of the windshield to watch the road ahead, while trying to listen to the voice coming from the phone by their ear (due, presumably, to the cross-modal links in spatial attention highlighted earlier). It is possible that performance in this situation could be improved if the speaker's voice were to be presented from directly in front of the driver (i.e., via a "talking windshield"), to take advantage of the underlying cross-modal links that constrain the deployment of endogenous spatial attention. Similarly, a better understanding of the cross-modal links in exogenous orienting may also lead to the design of more effective multisensory warning signals. Indeed, the latest research by Cristy Ho and her colleagues has shown that bimodal cues (i.e., audiovisual or audiotactile) appear to capture the spatial attention of car drivers far more effectively than unisensory warning signals (at least when the various unisensory cues are presented from the same spatial location, or direction).

Charles Spence

See also Attention: Divided; Attention: Selective; Attention: Spatial; Multimodal Interactions: Color–Chemical; Multimodal Interactions: Neural Basis; Multimodal Interactions: Pain–Touch; Multimodal Interactions: Tactile–Auditory; Multimodal Interactions: Visual–Auditory; Multimodal Interactions: Visual–Haptic

Further Readings

Ho, C., & Spence, C. (2008). *The multisensory driver: Implications for ergonomic car interface design.* Aldershot, UK: Ashgate.

Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Lawrence Erlbaum.

Spence, C., & Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention.* Oxford, UK: Oxford University Press.

ATTENTION: DISORDERS

Attention can imply various skills, such as staying on-task (as when driving down a monotonous highway), multitasking (as when talking to a passenger while driving), or focusing on one specific thing (as when reading a road sign). Students of perception are usually concerned with this last meaning of attention as a process of selection. Insights into how selective attention works can be gained from studying people in whom selection has broken down because of brain damage (usually caused by stroke). Attention is normally considered the gateway to conscious awareness, and disorders of attention thus entail losses of awareness for certain parts of the world. This entry discusses three major disorders of attention, and considers what sorts of perception may be possible for things outside of awareness.

Neglect

Neglect is a common and disabling consequence of brain damage. When it is severe, the eyes and head deviate toward the side of the damage, and people may fail to dress or groom the opposite half of their bodies, to eat food from that half of their plates, or to acknowledge anything on that side. Neglect is strongly associated with damage to the right side of the brain, particularly the posterior parietal and superior temporal lobes, so it is usually the left side of space that is neglected. The problem is not one of sensory transmission. The independence of neglect from sensory factors was illustrated elegantly by Eduardo Bisiach and Claudio Luzzatti, who asked two patients to describe a familiar Milanese square from memory. Both described buildings on the right, but not on the left, relative to their imagined viewpoints, showing that even mental images may be neglected.

Neglect is usually understood as a spatial bias of attention, such that items in relatively rightward positions are selected at the expense of those further to the left. Unawareness is not restricted to the visual world. Patients may neglect sounds, touches, and even smells. They may also neglect internal sensations, contributing to a loss of awareness for one side of the body. However, the world is experienced as complete because, by definition, the patient is unaware of the neglected parts. It may thus be difficult for a patient to achieve direct insight into his or her problem, and neglect can be hard to treat. At a theoretical level, the condition offers a unique window on spatial attention. No two patients are exactly alike in their symptoms, and the differences between them can inform us about how attention is organized in the brain. For instance, a patient who neglects visual objects might nonetheless be alert to touches on that side, or vice versa. Such patterns show that attention is not a single function shared across sensory modalities. The mechanisms by which we attend to our different senses must be at least partially separate because they can be disrupted separately by brain damage.

Research into neglect has tended to focus on its visual effects, but even here the symptoms are remarkably varied, suggesting many subdivisions of visual attention. For instance, visual space can be divided conceptually into that which can be accessed directly by reaching and grasping (near space) and that which cannot (far space). This distinction may be of significance for the control of behavior, and neglect research has helped confirm its biological reality. Peter Halligan and John Marshall reported a neglect patient who made large errors in estimating the midpoints of horizontal lines in near space, yet who could accurately bisect lines of equivalent visual extent in far space, by pointing a light-pen or throwing a dart. A further, distinct region of space is personal space, bounded by and immediately surrounding the body. Patients have been reported who can search both sides of external space effectively, yet fail to notice bright markers, such as balls of fluff, attached to one side of their clothing. Visual neglect for personal space may interact with neglect of bodily sensations to destroy awareness for one side of the body.

Visual neglect also highlights the fact that the apparently straightforward concepts of left and right are far from simple because their definitions depend on the spatial reference frame adopted. A primary distinction is drawn between egocentric and allocentric reference frames: the former specify positions relative to the viewer; the latter relative to things within the environment. If you look at a building, then tilt your head to one side, the building seems to stay upright despite your altered viewpoint. Your perception of up and down (and left and right) is not determined by your egocentric viewpoint alone, but also by environmental cues (in this case, the building itself, other objects in the scene, and gravity). Similar manipulations have been applied to patients with neglect, for instance asking them to search for objects in a display while sitting upright or lying on one side. Under such conditions, some objects may be neglected on the left side with respect to the current viewpoint (i.e., the bottom of the display if the patient is lying on his or her left side), and some on the left of the display regardless of posture. These patterns suggest that egocentric and allocentric factors interact to determine the space that is neglected.

One special class of allocentric reference frame is object-centered. If a familiar object (such as a building) has a recognizable top and bottom, we tend to think of it as having left-and-right sides that are independent of its positioning within the visual field. Consistent with this idea, neglect can affect the left side of individual objects, regardless of their position or orientation in space. For instance, a patient who usually neglects letters at the left end of words may continue to neglect the *initial* letters of words shown upside down, even though these are now seen on the right (e.g., reading PEAR as EAR). Patients may even show object-centered neglect at the same time as neglecting whole objects within the left part of space. These remarkable observations imply that selective attention uses spatial representations encoded with respect to multiple frames of reference. These representations are flexibly created, and re-created, as different frames of reference become relevant to the task at hand.

The diversity of symptoms across patients with neglect shows that it is not a single entity, but an umbrella term covering a constellation of related impairments. Moreover, in any given case, the clinical picture may depend on biased spatial attention and on co-occurring consequences of brain damage. These include reduced arousal and vigilance, which aggravate the lack of attention to the neglected side, and spatial memory problems, which impede structured searching of space. Neglect is thus a syndrome of many components, not all of which are necessary for neglect, but each of which can color its expression. At its core, however, is the skewing of attention away from one side. This may sometimes appear as a simple lack of attention for one half of space, but what is more typically seen is a directional bias across the entire space, with things in relatively rightward positions exerting a stronger pull on attention than do those that are relatively leftward.

Extinction

In the most basic assessment of visual extinction, the examiner faces the patient and holds up a finger on either side, while the patient looks straight ahead. The examiner then flexes one or both fingers rapidly, and the patient must report which fingers moved. Extinction is diagnosed if the patient reports single flexions on either side, but only one when both occur together: one event "extinguishes" awareness of the other. Analogous methods are used to diagnose extinction in other modalities (e.g., touch, hearing, smell). Like neglect, extinction reflects the outcome of a spatially biased competition for attention. However, it is elicited only when the stimulation is brief. Extinction may follow damage to either side of the brain, affecting the side of space opposite to the damage. Unlike neglect, extinction shows no preferential association with the right hemisphere. This has led some authors to suggest that it is separate from neglect, though others believe it to be a milder form of the same condition.

Morris Bender, who conducted the pioneering work on extinction, viewed it as a sensory disorder in which subtly impaired sensation on one side is exposed by stimulus competition. His observations of tactile extinction in patients with spinal injuries showed that the symptom can arise from a sensory imbalance. Nonetheless, numerous phenomena have been found that defy sensory accounts and imply an attentional disorder. For instance, tactile extinction can be determined by position in external space rather than by the sensory surface: if a touch to the right hand extinguishes awareness of a touch to the left, then this reverses when the hands are crossed over the midline. Similarly, visual extinction can occur when the competing stimuli are within the same half of visual space, showing that it depends on relative, rather than absolute, location. Visual extinction is also affected by various perceptual properties of the stimuli. For example, if two brackets are presented, extinction may be less likely when they face one another () than when they oppose one another)(. The former pair constitutes a better group according to Gestalt laws of perceptual organization, and are selected together rather than competing for attention.

Simultanagnosia

Simultanagnosia, one of the key features of Bálint's syndrome, typically requires damage to the posterior parietal lobes on both sides of the brain. The condition is reminiscent of extinction, in that the patient perceives only one object when multiple objects are present. However, whereas extinction requires brief stimulation and is determined by spatial location, simultanagnosia persists with prolonged viewing and is insensitive to the layout of a scene, being determined instead by the objects within it. A patient with simultanagnosia will report being aware of only one object at a time. The focus may change intermittently, so that several objects are eventually perceived, but there is no experience of their simultaneous presence. In describing a scene, such a patient may give a halting commentary, naming one object and then another, without grasping their interrelationships. A man posting a letter could be, "a man . . . a letter . . . a postbox."

The existence of simultanagnosia provides reason to believe that visual attention can select specific objects, rather than just regions of space. Simultanagnosia cannot simply reflect spatially restricted attention because the object selected can be of any size, from a single snowflake to Mount Everest. Moreover, if two line drawings are overlapped so that they occupy the same region of space, only one of them will be perceived. Being object-based, simultanagnosia can allow researchers to probe what constitutes an object in vision. The governing principles are again predicted well by the Gestalt laws of perceptual organization. Alexander Luria famously showed that a patient with simultanagnosia saw only one of two overlapping triangles of different colors, but a Star of David (✡) when the triangles shared the same color, thereby cohering as one object. Similarly, only one of two side-by-side circles was seen, but a pair of spectacles was reported when a horizontal line connected the circles.

The Fate of Unattended Things

Neglect, extinction, and simultanagnosia are disorders of attention, not of sensory transmission, so the unattended stimuli are potentially available in the brain for processing. Researchers may thus ask what aspects of perception can proceed without attention, and whether things that escape awareness can shape behavior. As noted, the likelihood of a stimulus being unattended is influenced by Gestalt grouping factors, suggesting that the perceptual organization of a scene into its constituent parts is accomplished before attentional selection. Many other studies have reported that more complex aspects of unattended stimuli, even identity and meaning, can affect behavior. In the most famous such experiment, a neglect patient was presented with two drawings, one of which had flames emerging from its left side. The patient maintained that the houses were identical, but when forced to choose which house she would prefer to live in, reliably tended to choose the one that was not on fire. She avoided the flames unconsciously, despite expressing no awareness of them. Under some circumstances, unattended things may even have as strong an influence as attended things. For instance, a patient with visual extinction was found to avoid an obstacle during reaching, to exactly the same extent, regardless of whether he could report the obstacle's existence. Disorders of attention are thus powerful "natural experiments" that allow us to query the purpose of conscious awareness, by studying what can be achieved without it. Sometimes, this is a surprisingly large amount.

Robert D. McIntosh

See also Attention: Object-Based; Attention: Selective; Attention: Spatial; Attention and Consciousness; Body Perception: Disorders; Consciousness: Disorders; Neuropsychology of Perception; Perceptual Organization: Vision; Perceptual Segregation; Unconscious Processes

Further Readings

Bisiach, E., & Luzzatti, C. (1978). Unilateral neglect of representational space. *Cortex, 14,* 129–133.

Driver, J., & Vuilleumier, P. (2001). Perceptual awareness and its loss in unilateral neglect and extinction. *Cognition, 79,* 39–88.

Halligan, P. W., & Marshall, J. C. (1991). Left neglect for near but not far space in man. *Nature, 350,* 498–500.

Heilman, K. M., Watson, R. T., & Valenstein, E. (2003). Neglect and related disorders. In K. M. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (4th ed., pp. 296–346). Oxford, UK: Oxford University Press.

Marshall, J. C., & Halligan, P. W. (1988). Blindsight and insight in visuospatial neglect. *Nature, 336,* 766–767.

McIntosh, R. D., McClements, K. I., Schindler, I., Cassidy, T. P., Birchall, D., & Milner, A. D. (2004). Avoidance of obstacles in the absence of visual awareness. *Proceedings of the Royal Society of London B, 271,* 15–20.

Robertson, L. C. (2004). *Space, objects, minds, and brains.* New York: Psychology Press.

ATTENTION: DIVIDED

For experienced drivers, talking on a cell phone while driving seems relatively easy. From a processing perspective, however, it requires the concurrent perception, analysis, decision making, and response selection of multiple stimuli from different sensory modalities. Research on the efficiency of such concurrent perceptual and cognitive processing of multiple stimuli or tasks comes under the heading of *divided attention.* The degree to which performance deteriorates under such conditions has been studied to characterize limitations on real-world human performance, as well as to inform theoretical models of the functional architecture of the human perceptual-cognitive system. Theories of divided attention have proposed that limitations on multitask performance can be attributed to competition for processing *resources* as well as to competition for processing *mechanisms.*

Laboratory research on divided attention typically involves dual-task methodology, in which performance on a given task performed alone is compared with performance when the same task is performed in combination with another task, whose difficulty is often systematically varied. Deterioration under dual-task conditions is assumed to reflect limitations on divided attention, whereas the lack of any decline in performance is assumed to reflect successful division of limited cognitive resources. For example, despite the apparent ease with which one can talk on a cell phone and drive, laboratory studies reveal significant impairment in driving performance with cell phone use. This is true even with hands-free phones, suggesting that attentional resources cannot be effectively divided between the two tasks.

Not all dual-task decrements necessarily reflect a failure of divided attention. For example, deterioration in certain combinations of tasks or stimuli may reflect structural limitations on sensory processing rather than attentional limits. Thus, the need to look at one's cell phone to dial a number may produce a decrement in concurrent driving performance simply because one's eyes are no longer "on the road." Such a decrement would not be considered an attentional limitation. This entry describes theories of divided attention and effects of practice on the ability to perform multiple tasks simultaneously.

Theories

Everyday experience suggests that the ability to divide attention among stimuli or tasks depends on the nature and difficulty of the constituent stimuli or tasks. For example, reading e-mail messages while listening to a speaker deliver a lecture is intuitively more difficult than is reading e-mail while listening to music. The results of laboratory studies using dual-task methodology confirm these intuitions; the efficiency with which multiple tasks are performed simultaneously depends crucially on the particular combination of tasks and their difficulty. Theoretical attempts to account for and predict the patterns of interactions between tasks fall into two broad classes: *capacity theories* and *structural theories*. These classes of theory are not mutually exclusive, and neither class is able to account for the full range of data, suggesting that a comprehensive theory of divided attention performance may require multiple constructs.

Capacity Theories

Emerging in the 1960s and refined in the 1970s, *capacity theory* is based on the assumption that mental tasks (perception, memory, language, reasoning, problem-solving, etc.) require the graded allocation of cognitive resources (sometimes referred to as mental fuel or effort), and that the available pool of resources to carry out these tasks is finite. Thus, if the resources required by a particular combination of stimuli or tasks exceeds the resources available, performance on one or both tasks will decline, depending on the voluntary allocation policy (i.e., which task is given priority). Note that according to this perspective, the cognitive architecture is such that mental tasks can be carried out at the same time, but the efficiency of task processing varies continuously with resource availability.

General Versus Specific Resources

In its original conception, capacity theory assumed that the resources available for cognitive processing are relatively nonspecific. That is, all mental tasks draw from the same pool of resources, and therefore interference between tasks should be determined only by the amount of resources required by the constituent tasks, rather than by the nature of the tasks or stimuli (e.g., their similarity). However, as noted earlier, and as revealed by subsequent research, the more similar the tasks, the more dual-task interference is generally observed.

One highly influential account of similarity effects in dual-task interference is the notion that there are a number of separate pools of resources that are specialized for particular types of cognitive processes. According to this account, interference will be observed only if the constituent tasks draw from the same specific pool of resources. If the tasks draw from different resource pools, then there should be no interference, even as the difficulty of the tasks is increased.

The most well-developed version of this account, *multiple resource theory,* proposes a set of resource pools defined across three dimensions: (1) stages of processing (perceptual-central vs. response), (2) format of internal representations (verbal vs. spatial), and (3) format of response (manual vs. vocal). Combinations of values across these dimensions are assumed to define a specific resource pool. Evidence consistent with this account comes from a host of dual-task studies that have systematically manipulated constituent tasks with respect to the proposed dimensions.

Despite the usefulness of this model, there are still cases where combining tasks that appear to share little with respect to processing dimensions nonetheless produces significant dual-task interference. For instance, even though cell phone use involves auditory input, verbal representations, and verbal output, it interferes with driving, which involves visual input, spatial representations, and manual output. Such examples would appear to indicate at least some degree of undifferentiated resource sharing.

Structural Theories

Structural models of divided attention performance emerged from the classic dichotic listening studies of the 1950s and 1960s, in which separate messages were presented to each ear. These experiments, which were modeled after early air traffic control environments, revealed severe limitations in people's ability to divide attention between the two messages. In particular, when attending to one message, only crude aspects of the unattended message were noticed (e.g., the gender of the voice delivering the message). In contrast to capacity theories, which attribute such limitations to competition for resources, structural theories implicate competition for processing *mechanisms* or *operations*. The notion is that critical mental operations can only be applied to one stimulus or task at a time, resulting in a processing bottleneck when attempting to simultaneously process multiple stimuli or tasks that all require that operation.

Perceptual Limitations

One proposed locus for a structural processing bottleneck is in the perceptual operations associated with stimulus identification or pattern recognition. According to this view, there is a limit on the number of stimuli that can be identified at one time. The typical experiment designed to evaluate this proposal involves measuring changes in performance as a function of the number of simultaneously presented stimuli. Generally, performance does indeed decline as the number of display items increases. For example, in visual search tasks where observers report the presence or absence of a prespecified target, both response times and error rates tend to increase with the number of distractors in the display, and this is particularly true when the distractors are perceptually similar or share critical properties with the target.

Central Limitations

In addition to limitations on perceptual processing, evidence also shows that central processes may also represent a locus for a structural bottleneck. In particular, research suggests that the process of choosing and selecting an appropriate response can only be applied to one task or stimulus at a time. Thus, when attempting to perform two tasks at the same time, processing on one of the two tasks will have to be suspended temporarily until response selection for the other task is complete, which necessarily produces a decline in task performance.

Evidence for this notion comes from experiments in which subjects must make speeded responses to one stimulus (Task 1) followed, at varying intervals, by a second stimulus (Task 2). For example, subjects might respond to the pitch of a tone (high or low) followed by the identity of a digit (1, 2, or 3). As the time interval between the two stimuli is decreased, performance on Task 2 typically slows down (assuming Task 1 is given priority). This phenomenon, known as the *psychological refractory period,* is consistent with a processing bottleneck, in that as the interval between the two stimuli decreases, and processing of the two tasks overlaps to a greater and greater extent, some aspect of Task 2 processing has to wait until response selection for Task 1 is complete.

Evidence that the bottleneck in these experiments consists of response selection comes from systematic manipulations of the difficulty of Task 2. Difficulty effects that influence processing before response selection (e.g., the difficulty of perceptual processing) are actually reduced as the amount of dual-task slowing increases. This makes sense if perceptual processing of difficult stimuli can continue during the time Task 2 is waiting for response selection of Task 1 to finish. In contrast, difficulty effects that influence response selection itself (e.g., the intuitiveness of the mapping between stimuli and responses) simply add on to dual-task slowing, indicating that response selection cannot occur during the wait time. In other words, response selection itself constitutes the processing bottleneck. However, there are exceptions to this pattern. In particular, studies have shown that with practice on the constituent tasks, or with particular combinations of stimulus and response modalities, response selection for the two tasks can proceed concurrently.

Effects of Practice

As discussed earlier, the difficulty of tasks has a strong influence on the ability to perform multiple tasks simultaneously. The difficulty of a given task is influenced by the type or amount of processing required, as well as by the amount of experience or practice one has had with the task. For example, although novice drivers may be unwilling to carry on a conversation with a passenger while driving, the situation changes dramatically with only moderate amounts of driving experience. With practice, tasks become subjectively easier. In the context of capacity theories, improvement in performance with practice is assumed to reflect a reduction in the amount of resources needed to accomplish the task. At the extreme, a highly practiced task may become "automatic" in the sense that few, if any, limited cognitive resources are required to accomplish the task. For example, studies of visual search have shown that when observers search for the same target letters for thousands of trials, the number of distractor letters present in the display has no effect on response times, suggesting that the identification of the targets no longer requires the controlled allocation of limited processing resources. It follows that as the resource demands of a particular task decline, the ability to perform that task in combination with other tasks should improve, and this prediction is indeed supported by laboratory research. In the context of structural theories, practice is assumed to reflect an enhanced ability to coordinate critical processing operations. For example, practice may improve the ability to rapidly switch the response selector back and forth between tasks.

Charles L. Folk

See also Attention: Effect of Breakdown; Attention: Effect on Perception; Attention: Selective; Cell Phones and Driver Distraction; Change Detection

Further Readings

Cherry, E. C. (1953). Some experiments on the recognition of speech with one and two ears. *Journal of the Acoustical Society of America, 25,* 975–979.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Meyer, D. E., & Kieras, D. E. (1999). Précis to a practical unified theory of cognition and action: Some lessons from the EPIC computational models of human multiple-task performance. In D. Gopher & A. Koriat (Eds.), *Attention and performance: Vol. 17. Cognitive regulation of performance: Interaction of theory and application* (pp. 17–88). Cambridge: MIT Press.

Pashler, H., & Johnston, J. C. (1998). Attentional limitations in dual-task performance. In H. Pashler, (Ed.), *Attention* (pp. 155–189). Hove, East Sussex, UK: Psychology Press.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review, 84,* 1–66.

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied, 9,* 23–52.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63–98). London: Academic Press.

ATTENTION: EFFECT OF BREAKDOWN

Many encyclopedia entries and introductory chapters written about attentional selection start by noting that our environment is rather rich in sensory stimulation that cannot possibly be processed simultaneously and there has to be a mechanism (attentional selection) that deals with this bombardment of sensory stimulation by processing only the relevant information. This statement is valid, but what is the evidence for this implied limited capacity (i.e., the breakdown of attention)? Perhaps you have had the experience of walking into a crowded café trying to find an empty table while holding a tray with hot soup and a sandwich. After quite a bit of searching around, you finally find a table and sit down to enjoy your food. But just as you are about to bite into your sandwich, a friend of yours comes by your table asking why you've ignored her waving at you. This example illustrates a simple yet startling fact—even though we may feel as if we perceive everything around us, the perceptual reality is that we are only aware of a small subset of sensory events at any single moment in time. The failure of attention, illustrated with this example, reflects the fact that most of the sensory information available in the physical world is actually not processed, and therefore is not available for conscious perception. This lack of conscious access to the sensory information present in the physical world is termed *breakdown of attention.* Such apparent breakdown, or limited capacity, arises because there is too much distraction from all the possible sensory input that competes for conscious perception. This entry focuses on instances that demonstrate the need for attentional selection, give examples of breakdowns of attentional selection, and briefly considers the fate of the information that failed to reach conscious perception.

Behavioral Phenomena

One of the most striking examples of the breakdown in the attentional system, caused by sensory overload, was demonstrated by Daniel Simmons and his colleagues. The authors conducted a series of experiments that revolved around a simple scenario—a young man (one of the experimenters) approaches a passerby (unsuspecting participant in this experiment) asking for directions on how to get to one of the buildings on campus. In the middle of the conversation, however, two people carrying a door (experimental accomplices) walk in between the experimenter and the participant briefly obstructing the participant's view of the experimenter. During the time that the experimenter was invisible to the subject, another person trades places with the experimenter, so that once the door has passed, the participant is left talking to a new person! This experiment found that only 50% of the participants noticed that, after the interruption, the identity of the person they were conversing with was switched, even though the two experimenters were dissimilar in build, had different haircuts, had different pitched voices, and wore dissimilar clothing. This example is particularly potent because the experiment was conducted not in the contrived laboratory environment, but rather in our natural environment (i.e., outside) with innocent bystanders serving as participants. In addition, this experiment was not conducted in the middle of a hustle and bustle of, say, Times Square (an environment that we often think of as being conducive to sensory overload). This experiment demonstrates a striking breakdown of attention in our day-to-day environment, even under conditions when part of the stimulus was attended (few would argue that two people who are carrying a conversation with one another do not, at least on some level, pay attention to each other's identity). The breakdown of attention in this experiment could be attributed to competition from a multitude of sensory and mental events related and unrelated to the actual task at hand of giving directions—for example, remembering the campus layout, remembering landmarks on the way, hearing birds chirping in the background, or feeling faint pain from a blister on the left foot. All these cognitive tasks place enormous constraints on attentional selection, ultimately resulting in a failure of perception.

Such failures of perception are the rule rather than an exception. As such, we live in a limited world of what our attentional selection mechanisms have reconstructed for us. An insight into just how limited our perceptual world is without attentional selection is gained from examining perceptual experiences of patients with hemispatial neglect. This disorder is termed *hemispatial neglect* ("neglect") to reflect the failure to attend to information appearing on the contralesional left side.

Recipe for Breakdown—Competition and Attentional Load

What are the circumstances that lead to breakdown of attention? Robert Desimone and John Duncan proposed a biased competition model for selective attention suggesting that items presented in a multielement scene are not processed independently, but rather interact in a mutually suppressive way (i.e., each items competes for processing, increasing the odds of attentional failure). In other words, processing a single item within the scene will not be enough to tax capacity limitations, whereas multiple items within the scene will compete for sensory processing and will compete for attentional resources. Such suppressive interactions have been observed not only behaviorally but also with neurophysiological and neuroimaging techniques. For example, Sabine Kastner and her colleagues examined the consequence of suppressive interactions (i.e., when multiple sensory stimuli compete for processing) using functional magnetic resonance imaging (fMRI). The authors presented participants with a set of four colorful pictures that were either presented simultaneously (i.e., simultaneous condition) or one at a time (i.e., sequential condition) and participants were simply asked to passively view the presented items. fMRI blood-oxygenated-level-dependent response (BOLD) was recorded during presentation of sequential condition (summed over presentation of four items) and during the simultaneous condition. The overall fMRI response was much reduced in the simultaneous condition compared with the sequential condition. This finding demonstrates that stimuli presented in close spatial proximity do compete for attentional resources, thus inhibiting and interfering with one another. Such competition is eliminated, however, when items are presented one at a time.

In a recent investigation, Sarah Shomstein and Steven Yantis demonstrated that competitive interactions are not limited to a single modality. Rather, competition is observed when multiple stimuli, of any kind, vie for attentional resources. In this study, participants were presented with visual and auditory information simultaneously and were asked to pay attention to one or the other depending on instructions, and to occasionally shift their attentional focus from visual stimulation to auditory stimulation (and vice versa). Using fMRI, the authors demonstrated that overall activity in the corresponding sensory regions (i.e., auditory cortex and visual cortex) was decreased in presence of competing sensory information, similar to the simultaneous condition in the Kastner study described earlier. Moreover, even though the subject was not attending to the competing information, it was observed that when participants were paying attention to the visual stimuli, activity in the visual cortex increased and activity in the auditory cortex decreased. The opposite pattern was observed when auditory information was attended. These findings support several conclusions. First, early sensory cortical responses, as measured by the fMRI signal in auditory and visual cortices, are modulated by attention. Second, the "push–pull" effect of switching attention between vision and audition suggests a neural basis for behavioral evidence that focusing attention on auditory input (e.g., a cellular telephone conversation) can impair the detection of important visual events (e.g., what's happening on the road while driving an automobile). When attention must be directed to audition, the strength of early cortical representations in the visual system are compromised (and vice versa), leading to potentially significant behavioral impairments.

Another contributing factor to breakdown of attentional selection, bearing directly onto biased competition, is attentional load. According to the load hypothesis, originally proposed by Nilli Lavie, the extent to which distracting elements compete for resources is a function of how attentionally demanding the task is. In other words, under high perceptual load (e.g., reduced size or contrast of stimuli or number of stimuli), attentional capacity will be maximally taxed, thus resulting in increased competition, but under conditions of low perceptual load, there would be enough capacity to enable processing of all the available sensory input, reducing the need for competitive interactions.

The Fate of the Unattended

Although much recent research has been devoted to understanding the psychological and neural processes underlying attentional selection, researchers know little about the fate of the sensory information that is not selected. In other words, the question is what happens to all the information that is not processed, that is not available for conscious perception? Did the breakdown of attention result in total and utter failure of perception, or is the unattended sensory information available for perception and for behavior in some capacity? For example, when an observer selectively attends to some subset of the visual input, the degree to which the remaining, unattended input is represented remains largely unknown. That our current knowledge of such scenarios is so underspecified may be surprising given that understanding the nature and extent of processing of unattended information is one of the long-standing and seminal issues in perception. It is not particularly contentious that after the light hits the retina, or the air moved by a sound excites hair cells in the cochlea, some amount of processing is accomplished, regardless of whether the specific information is attended or not. What remains more uncertain, however, is the extent to which the unattended information influences behavior.

A strong claim made by some researchers is that little, if any, visual processing can occur in the absence of attention and that perception cannot proceed without attentional selection. Evidence supporting this account comes from inattentional blindness paradigms, pioneered by Arien Mack and Irvin Rock, in which task-irrelevant grouped items went unnoticed and did not influence the behavioral performance of the observers. In one well-known version of this paradigm, observers judge whether the horizontal or vertical arm of a briefly flashed cross was longer. On the fourth trial, an unexpected item—a word, a face, a shape, a texture—was presented simultaneously with the cross but at a different position on the screen (within the arms of the cross, but not at the cross section of the arms). When the cross was presented at fixation and the unexpected item was presented parafoveally (peripherally), about one quarter of the observers failed to perceive the presence of the unexpected object. Even more startling, when the cross was presented parafoveally and the unexpected item was presented at fixation (i.e., precise location of where the participant was gazing directly), nearly three quarters of observers failed to detect the unexpected item. These findings suggest that, in the absence of attention, information is not processed and therefore perception fails.

An equally strong but opposing account suggests that fundamental visual processes such as perceptual grouping and figure-ground segmentation can take place in the absence of attention. In one illustrative study, Cathleen Moore and Howard Egeth asked participants to judge which of two parallel lines, superimposed on a background matrix of black-and-white dots, was longer. The background dots, whose presence was orthogonal to the line-judgment task, were either randomly colored or were grouped to form the Müller-Lyer illusion (optical illusion that consists of two sets of arrow-like figures, one with both ends pointing in, and the other with both ends pointing out; when asked to judge which of the two lines is longer, viewers typically chose the inward pointing arrow—see Cultural Effects on Visual Perception, Figure 1). Because the length judgment was clearly influenced by whether the background dots gave rise to the illusion, the authors concluded that background dots, though not attended, were in fact perceptually processed and as such affected behavior. Further support for the view that unattended information influences behavior comes from a recent study by Ruth Kimchi and Irene Razpurker-Apfeld, in which subjects performed a change detection task on a small, centrally located black-and-white matrix, presented for a brief duration. The matrix-like target stimulus was embedded in the center of a background pattern, consisting of colored dots that grouped by similarity into either rows, columns, or simple shapes. On each trial, two successive displays were presented, and subjects judged whether the central target matrix remained the same or changed. In addition, within a trial, the organization of the background elements stayed the same or changed, independent of the status of the target matrix. Of greatest relevance was the finding that the grouping of the background, unattended stimuli influenced the detection of changes in the target matrix, leading to the conclusion that unattended background elements were perceived to an extent that was enough to influence task-relevant perceptual judgments.

Sarah Shomstein

See also Attention: Effect on Perception; Attention: Object-based; Attention: Physiological; Attention: Selective; Attention: Spatial; Attention and Consciousness; Visual Search

Further Readings

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review Neuroscience, 18,* 193–222.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review Neuroscience, 23,* 315–341.

Kimchi, R., & Razpurker-Apfeld, I. (2004). Perceptual grouping and attention: Not all groupings are equal. *Psychonomic Bulletin and Review, 11*(4), 687–696.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *TRENDS in Cognitive Sciences, 9*(2), 75–82.

Mack, A., & Rock, I. (1998). *Inattentional blindness.* Cambridge: MIT Press.

Moore, C. M., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance, 23*(2), 339–352.

Shomstein, S., & Yantis, S. (2004). Control of attention shifts between vision and audition in human cortex. *Journal of Neuroscience, 24,* 10702–10706.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people in a real-world interaction. *Psychonomic Bulletin and Review, 5*(4), 644–649.

ATTENTION: EFFECT ON PERCEPTION

Attention is the mechanism that allows us to select relevant information for processing from the vast amount of stimuli we are confronted with, prioritizing some while ignoring others. Attention can affect perception by altering performance—how well we perform on a given task—or by altering appearance—our subjective experience of a stimulus or object. Visual attention is commonly divided into three types: (1) spatial attention, which can be overt, when an observer moves his or her eyes to a relevant location and the focus of attention coincides with the movement of the eyes, or covert, when attention is deployed to relevant locations without accompanying eye movements; (2) feature-based attention, which is deployed to specific features (e.g., color, orientation or motion direction) of objects in the environment, regardless of their location; (3) object-based attention (which will not be discussed further here). By focusing on different kinds of information, these types of attention optimize our visual system's performance: whereas spatial attention guides an observer to a particular location, feature-based attention guides an observer to a particular feature of the object or stimulus. For instance, when waiting to meet a friend at a café, we may deploy spatial attention to the door (where he or she is likely to appear) and feature-based attention to orange objects (because he or she often wears an orange jacket).

This entry deals with the effects of covert attention on performance and appearance. Note that overt attention also plays a pivotal role in selectively processing information. The eyes can be moved quickly and efficiently, enabling overt attention to compensate for the rapid decline of visual capacities away from the fovea. Thus, the perceptual consequences of overt attention at the attended location correspond to the immediate benefit of high-resolution foveal vision.

Spatial Attention

To investigate covert attention, it is necessary to make sure that observers maintain fixation, and to keep both the task and stimuli constant across conditions while manipulating attention. It is well established that covert attention improves perceptual performance—accuracy and speed—on many detection, discrimination, and localization tasks. Moreover, attention affects the appearance of objects in several tasks mediated by dimensions of early vision, such as contrast and spatial frequency.

There are two systems of covert attention: endogenous and exogenous. *Endogenous attention* refers to the voluntary, sustained directing of attention to a location in the visual field. Experimentally, a central cue—typically an arrow or a bar at fixation—points to the most likely location of the subsequent target. Within about 300 milliseconds (ms), attention can be deployed to that location and can be sustained there. *Exogenous attention* refers to the automatic, transient orienting of attention to a location in the visual field, brought about by a peripheral cue or a sudden abrupt onset of a stimulus at that location. Exogenous attention has a transient effect; its effectiveness peaks at about 80 to 120 ms and decays shortly thereafter.

Spatial Attention and Performance

Endogenous (voluntary) and exogenous (involuntary) attention have some common perceptual effects. Detection or discrimination of a target are typically better (faster, more accurate or both) in trials in which the target appears at the cued location than at uncued locations. For instance, both exogenous and endogenous attention improve contrast sensitivity and spatial resolution. Many studies have shown that spatial attention lowers the contrast threshold at which observers attain a given performance level when they perform tasks mediated by contrast sensitivity. Enhanced contrast sensitivity enables people to better detect the presence or absence of a stimulus, discriminate its characteristics (e.g., was the stimulus tilted left or right? moving upward or downward?), and localize it (e.g., was the stimulus in the right or left visual field?). Interestingly, whereas contrast sensitivity is enhanced at the cued location, it is decreased for the stimuli appearing at uncued locations. This is the case even when few clearly visible stimuli appear in the display, and observers know with certainty the stimuli locations and which stimulus to respond to.

Likewise, when spatial attention is directed to a given location, performance improves in texture segmentation tasks and acuity tasks that are limited by spatial resolution. For instance, when attending to a location observers can resolve information that is unresolvable without attending to that location and can discriminate finer details than they can without directing attention to the cued location. As is the case with contrast sensitivity, the improved performance at the cued location is accompanied by impaired performance at the uncued locations. The same pattern of results has been reported for human and nonhuman primates in acuity tasks.

Despite their similarities, endogenous and exogenous attention can yield different perceptual effects: Whereas endogenous attention improves perception, exogenous attention can in some conditions actually impair perception. For instance, in a texture segmentation task in which performance is influenced by spatial resolution, cueing attention improves texture segmentation in the periphery, where spatial resolution is too low for the task, but impairs performance at central locations of the visual field, where spatial resolution is already too high for the task. In contrast, endogenous attention improves performance across the visual field. Exogenous attention also impairs temporal order judgment, whereas endogenous attention improves it. These findings illustrate that both endogenous and exogenous attention affect performance in spatial (texture segmentation) and temporal (temporal order judgment) tasks, but that the mechanisms underlying these attention systems differ, with endogenous attention being more flexible.

Spatial Attention and Appearance

It is well established that attention improves performance on many tasks. However, for more than 100 years, psychologists, philosophers, and physiologists (e.g., Ernst Mach, Gustav Fechner, Hermann von Helmholtz, Wilhelm Wundt, and William James) have debated whether attention changes one's subjective experience of the visual world. Although the effects of attention on performance may suggest that attention affects the stimulus representation and subjective appearance, many authors have attributed these effects to a decision-making process or selective read-out. The question of whether attention alters appearance has rarely been investigated, though. Which aspects of our visual experience does attention affect? Can attention make a visual pattern seem more detailed, or a color more vivid? Much of the early work on this topic, which is relevant to the topic of subjective experience and awareness, was introspective and conflicting conclusions were often drawn from such subjective methods of investigation. Whether and how attention affects appearance has been systematically investigated only recently. This may be because of the difficulty in objectively testing and quantifying the subjective experience of perceived stimuli and changes in such experience with attention. It is important to distinguish a change in appearance from any bias that may arise because of experimental design and task demands.

A psychophysical paradigm developed by Marisa Carrasco and colleagues enables the assessment of the phenomenological correlate of exogenous (involuntary) attention, and makes it possible to study subjective experience and visual awareness more objectively and rigorously. This paradigm allows for simultaneous measurement of the effect of attention on appearance and performance. It manipulates exogenous attention via an uninformative peripheral cue and quantifies the observer's subjective perception using a task contingent on a comparative judgment between two stimuli with regard to a particular feature. For instance, to investigate the effects of attention on perceived contrast, observers are presented with two stimuli (Gabor patches, commonly used to investigate spatial vision; Figure 1a), one to the left and one to the right of fixation. Observers are asked to report the orientation of the higher contrast stimulus. These instructions emphasize the orientation judgment, when the main interest is in contrast judgments. Observers are not asked to directly rate their subjective experience on contrast, but to make a decision about a stimulus property, its orientation.

On each trial of this paradigm, the "Standard" stimulus is of a fixed contrast, whereas the contrast of the "Test" stimulus is randomly chosen from a range of contrasts near the standard contrast. The orientation of each stimulus is also chosen randomly. By flashing a dot (peripheral cue) briefly, exogenous (involuntary) covert attention is automatically directed to the cued location (at which either the Standard or the Test stimulus will appear).

Carrasco and colleagues have shown that by assessing observers' responses, it was possible to determine the contrast of the test stimulus that the observer judged to have the same contrast as the standard stimulus. This contrast was measured under three conditions: when the dot preceded (1) the location of the test stimulus, (2) the location of

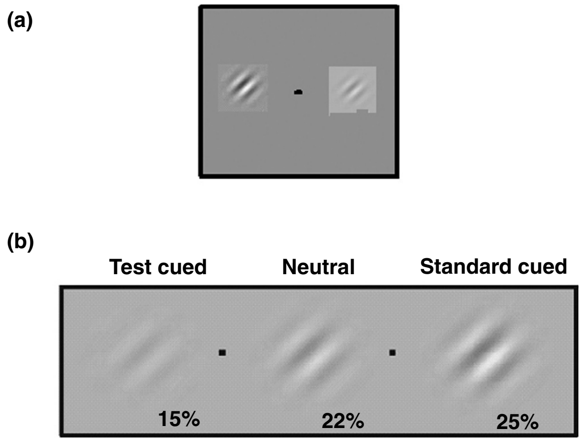


Figure 1 Effect of Attention on Contrast

Source: Carrasco, Ling, & Read, 2004.

Notes: (a) Display screen of the type of stimuli observers were shown by Marisa Carrasco and colleagues. Observers saw two stimuli, preceded by a brief neutral (at fixation) or peripheral (attentional) cue (appearing above one of the upcoming stimulus locations). The peripheral cue had equal probability of appearing on the left or right side, and was not predictive of the stimulus contrast or orientation. Observers were asked to indicate the orientation (left versus right) of the higher contrast stimulus. (b) Effect of covert attention on apparent contrast. If you were to look at 1 of the 2 fixation points (black dot), and the stimulus to the left of that fixation point was peripherally cued, the stimuli at both sides of that fixation point would appear to have the same contrast. With attention, when a 16% contrast stimulus is cued, it appears as if it had 22% contrast, and a cued 22% contrast stimulus appears as if it had 28%. (Note that this effect cannot be appreciated by inspecting the figure because a comparison between both patterns would lead to distributed allocation of attention to both patterns.)

the standard stimulus, or (3) the central location (fixation), so that covert focal attention was automatically directed toward the test stimulus, the standard stimulus, or neither stimulus. Exogenous attention was shown to significantly increase perceived contrast (Figure 1b). When observers' attention was drawn to a stimulus location, observers reported that stimulus as being higher in contrast than it actually was, thus indicating that attention changes appearance.

By coupling this paradigm with control experiments, cue bias or response bias have been ruled out as responsible for the observed effects. For instance, because of the ephemeral nature of transient attention (approximately 120 ms), lengthening the interval between the cue and target should eliminate any effect that it may have on perception, and any residual effect would be attributed to a cue bias. When the cue preceded the stimuli by 500 ms, neutral and peripheral conditions did not differ.

Most studies using this paradigm show that exogenous attention alters appearance of basic spatial (contrast, spatial resolution, color saturation, object size) and temporal (flicker rate, motion coherence, motion speed) visual dimensions. A modification of this experimental paradigm has revealed that endogenous (voluntary) attention also enhances perceived contrast. Although for many visual dimensions the attention effect on appearance is accompanied by an effect on performance (e.g., a higher proportion of correct responses in orientation discrimination), the performance effect is not necessarily mediated by the subjective change in appearance; for example, attention improves orientation discrimination of stimuli defined by hue, but it does not alter hue appearance.

By showing that the spatial deployment of attention leads to a change in phenomenological experience, these studies confirm that covert attention can intensify the perceptual impression of a stimulus. Attention affects how well we perform a visual task and it affects what we see and experience.

Feature-Based Attention

Feature-based attention is the ability to enhance the representation of image components throughout the visual field that possess a particular feature. Directing feature-based attention to specific features—such as color, orientation, and motion direction—increases performance for detecting, discriminating, or localizing those features across the visual field. This type of attention is critical when human or nonhuman primates search a display to detect a target among distracters.

Studies of feature-based attention generally control for spatial selection by using compound stimuli containing multiple features superimposed over the same spatial location, and requiring that observers attend to one of those features. For instance, in a compound motion stimulus (consisting of dots moving right and dots moving left), attending to one motion direction (e.g., dots moving right) produces a motion aftereffect consistent with the attended direction, and in a compound orientation stimulus (consisting of right-tilted and left-tilted lines), attending to one orientation (right-tilted lines) produces an orientation aftereffect consistent with the attended orientation. Psychophysical and neuroimaging studies have shown that the effects of feature-based attention occur at the spatially attended location as well as at non-attended locations across the visual field.

Marisa Carrasco

See also Attention: Covert; Attention: Object-Based; Attention: Physiological; Attention: Selective; Attention: Spatial; Attention: Theories of; Eye Movements: Behavioral; Psychophysical Approach; Visual Search

Further Readings

Carrasco, M. (2009). Visual attention alters appearance: Psychophysical studies of subjective experience. In T. Bayne, A. Cleeremans, & P. Wilken (Eds.), *Oxford companion to consciousness*. New York: Oxford University Press.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience, 7,* 308–313.

Itti, L., Rees, G., & Tsotsos, J. (Eds.). (2005). *The encyclopedia of the neurobiology of attention.* Amsterdam: Elsevier.

Kinchla, R. A. (1992). Attention. *Annual Review of Psychology, 43,* 711–742.

Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology, 52,* 629–651.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

ATTENTION: OBJECT-BASED

Efficiently representing visual information requires that the observer select only a fraction of the multitude of information that is available to the visual system at any one instant in time. Attentional selection is the mechanism by which a subset of incoming information is extracted from the complex sensory environment. The selection can be achieved in a top-down, goal-directed fashion (e.g., when one sets out to find one's keys on the kitchen counter), or in a bottom-up fashion in response to highly salient or novel stimuli in the input (e.g., a red strawberry among greener strawberries on a bush captures the viewer's attention). Early models of attentional selection suggested that attention is directed to particular regions or locations in space (the region of the keys or the region of the red strawberry), enhancing perceptual processing of information appearing in the attended location, whereas more recent models have also included object-based representations as possible candidates of attentional selection (i.e., the keys or the red strawberry as objects of attention themselves). This entry considers this latter, object-based form of attentional selection in which attention is guided or influenced by object structure.

Behavioral Phenomena

The primary signature associated with object-based attention is the enhanced processing of information belonging to or appearing within the confines of one object that is selectively attended. Numerous behavioral studies have confirmed this *single object advantage*. For example, in a now-classic study demonstrating the advantage in processing features from a single object conducted by John Duncan in 1984 (see Figure 1a), subjects were shown displays consisting of an outline box upon which a diagonal line was superimposed, thereby occupying roughly the same spatial region as the box (so that spatial attention could not enhance just one of the two "objects"). Subjects were required to make judgments about two features that appeared on the same object (e.g., line orientation and texture from the line, or box size and gap side from the box) or, in a second condition, when one feature appeared on each of the two different objects (e.g., line orientation and box size). The critical result was that subjects were less accurate in reporting two features when they came from the two different objects (one feature from each) than when they came from a single object. Indeed, under simple conditions, subjects identified two properties of a single object just as accurately as they identified one feature. These results were taken as favoring a view that the visual field can be segmented or parsed into separate objects and that attention can then be directed selectively to a single object, thereby facilitating the processing of all of its features.

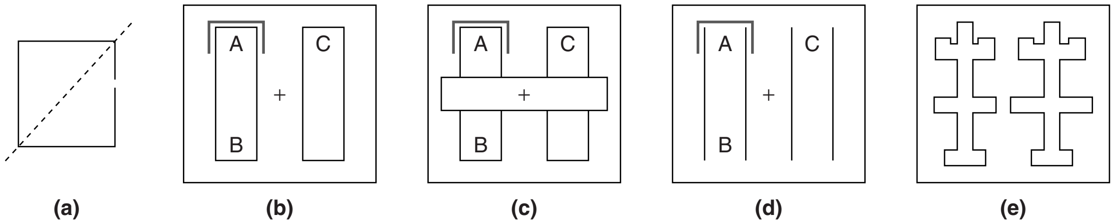


Figure 1 Examples of Object-Based Attention Displays

The single object advantage was also demonstrated in a further seminal study by Robert Egly and colleagues in 1994. In this study, subjects were presented with a display of two horizontally or vertically oriented rectangles. A cue appeared at one of the four ends indicating the most likely location in which the target would appear (see Figure 1b). Participants were required to detect the presence of the subsequently presented target (small white square). Unsurprisingly, detection times for the target in the cued location (valid trials; in Figure 1b with the cue in the upper left, this is the A position) were faster than for targets in the uncued locations (positions B or C). Most relevant for the current purpose is that reaction times were faster for targets that appeared at the uncued end of the cued rectangle (invalid; within object; position B) than for targets that appeared in the uncued rectangle (invalid; between object; position C), notwithstanding the fact that the spatial distance between the cue and both uncued locations was equivalent. The advantage for the uncued, within object location (position B) over the uncued, between object location (position C) is taken as evidence that all locations within an attended object are enhanced, the benchmark indicator of object-based attention, independent of spatial distance from the cue.

Several other behavioral paradigms have also successfully uncovered object-based attention. One such paradigm requires participants to compare two targets both of which lie on the same or on different objects—better performance is observed in the former than in the latter case. A further example is a variant of the Eriksen “flanker” task in which participants judge the identity of a target in the presence of flanking distractors that lie on the same or different objects: greater interference is observed when the flankers were in the same object, a consequence of enhancement of information within a single object. Taken together, such studies provide robust evidence that humans can select one object and, subsequently, all its information is preferentially facilitated.

Underlying Mechanisms

Having established the existence of object-based attention, much recent research has become increasingly concerned with the type(s) of visual representation and underlying mechanisms that constrain attentional processes. Indeed, growing consensus is that object-based attention may not be influenced only by object representations per se but also by the perceptual organization processes that the visual system uses to derive structure from the input. Consistent with this, many studies have uncovered object-based attention in displays in which the input can be organized into structured elements that adhere to the gestalt principles of grouping: thus, a single-object advantage can be observed for two spatially disparate parts of a single object that are occluded by a central bar (Figure 1c), for parts of an object that are discontinuous but benefit from shared symmetry, and for parts of an object that are separated but share common motion or color. Similarly, a cueing advantage has been reported for distant stimuli situated within arrays of evenly spaced parallel lines even when these lines do not form closed, coherent objects (see Figure 1d), as revealed by Alex Marino and Brian Scholl. Subparts of these holistically structured but disparate elements, which are perceptually organized, can also be selected, presumably by the same object-based attention process, but perhaps operating at a smaller scale, at the level of parts of an object. For example, this part-based attention was demonstrated in a study by Shaun Vecera and colleagues that borrowed the procedure from Duncan in which participants reported two attributes that appeared on the same part or on different parts of a single multipart object (see Figure 1e). Subjects reported two attributes in the display (e.g., gap side and arm length), and these could be from the same part or from two different parts of the same object. The report was more accurate when the attributes belonged to the same part than to different parts of the same object. Moreover, this part-based effect was not influenced by the spatial distance between the parts, ruling out a simple spatial attention interpretation of the results.

Several hypotheses have been put forward to account for the single-object attention advantage. One hypothesis maintains that object-based attention arises from the spread of attention (object-guided spatial selection)—when spatial attention is directed to one part of an object, there is a facilitation of early sensory processing that spreads to encompass other regions within the object's boundaries, hence the enhancement of all features and parts of the single object. This can be contrasted with the hypothesis that locations within the attended object are given higher priority for target search. If a target is not found at the cued location, other locations within the object's boundaries are searched before locations on uncued objects are searched. This attentional prioritization mechanism, which invokes a more top-down strategy, is proposed to control the order of locations to be searched, and on this account, the enhancement is not by virtue of early sensory processing. Ongoing investigations compare and contrast these views (as well as others) and much remains to be done to adjudicate between the alternative competing hypotheses.

Neural Basis of Object-Based Attention

Several recent investigations have begun to explore the neural mechanisms that underlie object-based attention and a series of functional magnetic resonance imaging (fMRI) and event-related potential (ERP) studies have been undertaken. One study, done by Notger Müller and Andreas Kleinschmidt using the Egly and colleagues paradigm described earlier, showed that the same object advantage is observable even in the early visual cortex (V1–V2). In an ERP study using a paradigm similar to that of Egly and colleagues in which a peripheral cue directed attention to the probable target location, an enhanced N1 component was documented at 130 to 150 milliseconds (ms) in response to targets in the cued or valid case, reflecting spatial attention, as shown by Antigona Martinez and colleagues. In addition, there was increased amplitude of the posterior, sensory-evoked N1 component over the interval 150 to 190 ms in response to uncued stimuli included within the common object, presumably reflecting the object-based selection. This site also serves as the source of these two signals, but one is earlier than the other, indicating that some of the same mechanisms likely support spatial and object-based forms of attention but that they may operate over different time scales.

Several studies (e.g., Liu et al.; Serences et al.; Shomstein & Yantis; Yantis et al.) have also explored the neural mechanism associated with the signal to switch attention between objects (i.e., the control system that triggers the switch between objects). Using fMRI, these studies have shown that the blood-oxygenation-level-dependent (BOLD) activity in the posterior parietal cortex is enhanced following instructions to shift versus to hold spatial attention, suggesting that a signal triggers the system to shift attention from one object to another. Shifting from one spatial position to another may engage different cortical regions than does shifting from one object to another; specifically, enhanced activity for the former engages the parietal cortex bilaterally, whereas shifting or redirecting attention between objects revealed activation in the left posterior parietal cortex. Interestingly, in this same study, object-sensitive shifts of attention were accompanied by modulations within the extrastriate regions of the occipital cortex (as in the study by Martinez et al. using ERP). These neuroimaging results elucidate the neural mechanism underlying object-based attentional selection described in many behavioral studies and reveal a network in which the consequences of attentional shifts triggered in the parietal cortex are manifest in earlier, sensory regions of the visual system. This dynamic circuit between the parietal and earlier visual regions presumably enables observers to focus preferentially on objects of interest that appear in complex visual scenes.

A final ongoing question concerns the relationship between the mechanisms mediating spatial and object-based attention and the extent to which these mechanisms are shared or separate. Specifically, some researchers have suggested that similar mechanisms might be at play; for example, some studies indicate that attention is spatially spread and then constrained by the object boundaries and groupings, and that this is true over a certain spatial extent. Moreover, as stated earlier, evidence from some fMRI and ERP studies (e.g., Müller & Kleinschmidt; Shomstein & Behrmann) report that similar brain areas are activated by spatial and object-based attention and that attention-based modulation of "within" object structure is evident even in early retinotopic areas of the cortex. This provides physiological evidence that directing spatial attention to one part of an object facilitates processing of the entire object, and thus, both spatial and object-based attention contribute to the single object advantage.

Marlene Behrmann and Sarah Shomstein

See also Attention: Spatial; Gestalt Approach; Visual Processing: Primary Visual Cortex

Further Readings

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*(4), 501–517.

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123*(2), 161–177.

Liu, T., Slotnick, S. D., Serences, J. T., & Yantis, S. (2003). Cortical mechanisms of feature-based attentional control. *Cerebral Cortex, 13*(12), 1334–1343.

Marino, A. C., & Scholl, B. J. (2005). The role of closure in defining the "objects" of object-based attention. *Perception & Psychophysics, 67*(7), 1140–1149.

Martinez, A., Teder-Salejarvi, W., Vazquez, M., Molholm, S., Foxe, J. J., Javitt, D. C., et al. (2006). Objects are highlighted by spatial attention. *Journal of Cognitive Neuroscience, 18*(2), 298–310.

Müller, N. G., & Kleinschmidt, A. (2003). Dynamic interaction of object- and space-based attention in retinotopic visual areas. *Journal of Neuroscience, 23*(30), 9812–9816.

Serences, J. T., Schwarzbach, J., Courtney, S. M., Golay, X., & Yantis, S. (2004). Control of object-based attention in human cortex. *Cerebral Cortex, 14*(12), 1346–1357.

Shomstein, S., & Behrmann, M. (2006). Cortical systems mediating visual attention to both objects and spatial locations. *Proceedings of the National Academy of Sciences USA, 103*(30), 11387–11392.

Shomstein, S., & Yantis, S. (2004). Control of attention shifts between vision and audition in human cortex. *Journal of Neuroscience, 24*(47), 10702–10706.

Vecera, S., Behrmann, M., & Filapek, V. (2001). Attending to the parts of a single object: Part-based selection mechanisms. *Perception and Psychophysics, 63*(2), 308–321.

Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., et al. (2002). Transient neural activity in human parietal cortex during spatial attention shifts. *Nature Neuroscience, 5*(10), 995–1002.

ATTENTION: PHYSIOLOGICAL

Often, when we're interested in an object, we move our eyes toward it, allowing us to see it more clearly. However, it has long been known that even in the absence of overt movements, perception of an object in the periphery can be improved by covertly shifting attention to it. *Covert attention* refers to our ability to internally modulate sensory processing of a selected stimulus. Attention can improve perception of the attended object, even though the input to the system (the physical stimulus reaching the retina, for example) is unchanged.

When we shift attention to an object, it may appear clearer, more distinct, or more intense, whereas ignored objects may seem less distracting or even fade from awareness. These perceptual effects are caused by changes in the responses of sensory neurons. Thus, one important issue in the physiology of attention is to understand the *effects of attention on sensory processing*. A second major issue is to understand how the brain controls which particular objects are attended. This is referred to as the *neural control of attention.*

Effects of attention on sensory neurons have been observed in all the major sensory systems. Most studies have found that attention increases the gain of sensory neurons encoding the attended object or feature, effectively amplifying signals from the attended stimulus. The effects of attention on sensory neurons become progressively stronger in higher-level brain areas compared with areas closer to the sensory input. Attention has been studied most extensively in the visual system, and effects of attention on visual neurons can be observed as early as the lateral geniculate nucleus (LGN), an area that receives input directly from the retina.

The control of attention has been studied more extensively for spatial attention than for feature attention. *Spatial attention* refers to the allocation of attention to specific objects or locations in the environment, whereas *feature attention* refers to the allocation of attention to particular sensory features (for example, the color red) regardless of location. Behavioral studies have shown that spatial attention can be attracted to objects or locations, either by virtue of their physical salience (for example, a bright flash of lightning tends to draw attention) or their behavioral relevance (for example, when driving, one can intentionally ignore a distracting billboard along the side of the road to attend to the cars ahead).

Physiological research suggests that spatial attention is controlled by a network of brain areas that includes the frontal and parietal cortex, as well as subcortical areas. Neurons in most of these areas are organized into maps that represent locations in the environment. The level of neural activity at a given location in the map is thought to represent the physical salience and behavioral relevance of stimuli at the corresponding location in the environment. The cells on the map having the greatest activity correspond to the most important location in the environment, and the greatest amount of attention is directed to this location. The assumption is that these maps influence sensory processing by modulating the activity of sensory neurons. Indeed, feedback connections linking attention control areas to sensory areas are known to exist and could serve as a conduit for these control signals to modulate sensory neurons. This entry describes the effects of attention on sensory neuron responses and neural control of attention.

Effects of Attention on Sensory Neuron Responses

Although the effects of attention on sensory neurons have been investigated most extensively for the visual system, it has been found across all the sensory modalities that higher-level brain areas show larger effects of attention than do areas located earlier in the processing stream. However, even within a given brain area, there is typically a great deal of variability in the extent to which attention influences the responses of individual sensory neurons.

Amplification of Signals Generated by Attended Stimuli

When attention is focused on a stimulus, sensory neurons encoding that stimulus tend to increase their firing rates, resulting in an amplification of the neural response to the attended stimulus. This amplification is similar in some ways to the increase in neural activity that would occur if the stimulus were more intense (e.g., higher contrast or louder). An amplification of neural responses is seen for feature attention as well as for spatial attention. For example, if attention is directed to downward motion, the gain of sensory neurons coding downward motion increases, and the gain of neurons coding upward motion tends to decrease.

Attention to a location or feature can cause sensory neurons to increase their activity even in the absence of an immediate stimulus. For example, if subjects are expecting an important object to appear at a certain location in the visual field, neurons that respond to that location may increase their baseline level of activity in anticipation of the appearance of the stimulus. Similar effects on baseline activity have been observed in the auditory system.

Effects of Attention on Encoding of Multiple Stimuli

The amplification of neural responses to an attended stimulus can have important implications when multiple stimuli are present in the environment. For example, in the visual system, each sensory neuron responds to visual stimulation within a restricted region of the visual field, called its *receptive field.* In complex visual scenes, two (or more) objects may lie within a single neuron's receptive field. A study of the visual system of monkeys by Jeffrey Moran and Robert Desimone found that when two stimuli are in the receptive field of a neuron, the neuron will predominantly respond to the stimulus to which the monkey is attending, rather than to an unattended stimulus. Thus, amplification of signals from the attended stimulus appears to reduce the influence of unattended stimuli on the responses of sensory neurons. Indeed, it is as though the receptive fields of sensory neurons effectively constrict around the attended item.

Effects of Attention on the Tuning of Individual Neurons

In addition to amplifying attended stimuli, one could imagine that attention might also sharpen the extent to which sensory neurons discriminate small differences in the features of an attended stimulus. For example, a visual neuron sensitive to the movement of objects in the visual field might be expected to discriminate smaller differences in the motion of an attended stimulus. However, with the exception of the sharpening of location tuning described previously, a sharpening of tuning for features has rarely been observed for individual sensory neurons. Thus, the primary mechanism of attention at the level of single neurons appears to be amplification of neural responses to the attended stimulus.

Synchronized Firing as a Mechanism of Attention

Focusing attention on an object changes the overall firing rate of neurons and can cause groups of sensory neurons encoding the attended object to fire in synchrony, such that the timing of their spikes is more closely matched in time. When neurons fire in synchrony, their activity tends to have a stronger effect on the downstream cells that receive their signals. An analogy would be to consider how well a message would be heard if a group of people all spoke it together in unison rather than speaking at different times. Thus, increased synchrony of sensory neurons could contribute to the improvements in perception that are afforded by attention.

Neural Control of Attention

The brain areas controlling attention have been studied primarily for the visual system. The posterior parietal cortex, prefrontal cortex, superior colliculus (located in the midbrain), and pulvinar nucleus of the thalamus are all believed to contribute to the control of visual attention. For example, neurologists have found that patients with damage to the right parietal lobe often have deficits in attending to objects on the left side of space. Even though these patients are not blind to stimuli in their left visual field, they tend to ignore people, objects, and events on that side, especially when a distracting stimulus is present in the opposite visual field. Damage to the frontal cortex in the vicinity of the frontal eye field can result in similar, although less severe deficits.

Physiological studies in monkeys have further delineated and refined our knowledge of the role of these areas in controlling attention. For example, it has been shown that if neurons in the parietal cortex, the frontal eye field, or the superior colliculus are activated, this activity can cause a shift of attention into the region of space encoded by the activated neurons. Other experiments have used microinjection of drugs to cause a temporary decrease in activity in the frontal eye field or parietal cortex. This decrease in activity causes temporary deficits in visual attention at locations in the visual field corresponding to the injected brain regions. These experiments provide direct evidence that neural activity in these structures controls the allocation of spatial attention.

Posner's Disengage-Shift-Engage Model

An early model for the control of attention was introduced by Michael Posner, who proposed that attention shifts could be decomposed into three elementary operations: first, attention must be disengaged from wherever it is currently focused; second, attention must be shifted to a new location; and, finally, attention must be engaged at this new location. On the basis of data from neurological patients, Posner proposed that the posterior parietal cortex was responsible for the disengage operation, that the superior colliculus was responsible for the move operation, and that the pulvinar was responsible for the engage operation.

A Salience Map for Shifting Attention

More recent models of attention shifting have focused on the concept of a salience map rather than on discrete disengage-shift-engage operations. A salience map is a topographical representation of external space that is laid out across the surface of a neural structure, such as the frontal eye field, posterior parietal areas, or superior colliculus. The activity of the neurons composing this map is affected by both the physical salience and the behavioral relevance of stimuli in the environment. Under this view, bottom-up signals related to physical salience (such as brightness or color) are sent from sensory areas to neurons in the salience map. The activity of neurons in the salience map is also strongly modulated by top-down signals related to the behavioral relevance of stimuli, which depends on the internal goals of the animal. Attention is allocated to the location in space corresponding to the peak of activity across this map. Specifically, signals from the salience map are believed to modulate the activity of neurons in sensory areas, such that sensory responses to stimuli at locations with strong activity in the map are boosted, as described in the previous section. In this way, perception of the most important stimuli is improved.

Robert M. McPeek

See also Attention: Covert; Attention: Disorders; Attention: Effect on Perception; Attention: Spatial

Further Readings

Chapman, C. E., & Meftah, el-M. (2005). Independent controls of attentional influences in primary and secondary somatosensory cortex. *Journal of Neurophysiology, 94,* 4094–4107.

Goldberg, M. E., Bisley, J. W., Powell, K. D., & Gottlieb, J. (2006). Saccades, salience and attention: The role of the lateral intraparietal area in visual behavior. *Progress in Brain Research, 155,* 157–175.

Hromádka, T., & Zador, A. M. (2007). Toward the mechanisms of auditory attention. *Hearing Research, 229,* 180–185.

Maunsell, J. H. R., & Cook, E. P. (2002). The role of attention in visual processing. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 357,* 1063–1072.

Moore, T. (2006). The neurobiology of visual attention: Finding sources. *Current Opinion in Neurobiology, 16,* 159–165.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229,* 782–784.

Posner, M. I., Cohen, Y., & Rafal, R. D. (1982). Neural systems control of spatial orienting. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 298,* 187–198.

Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annual Review of Neuroscience, 27,* 611–647.

ATTENTION: SELECTIVE

We experience the world around us, in all its sensory detail and complexity, directly—or at least so we believe. Yet a simple test reveals that our sense of direct experience may be misleading. If you are asked to reflect, with your eyes closed, on the details of the picture on the wall, or the number of cars on the road, or the shape of a stranger's eyes, the limits in processing information become evident. At times, we may not even notice when a significant part of a visual scene has been removed or altered, seemingly in front of our eyes. Attention is the portal through which sensory information is selected for more detailed examination, classification, and registration. *Selective attention* refers to selection of certain sensory inputs for privileged perceptual processing, further analysis, and improved transfer to memory. Attention selects some regions or objects in the visual display, or some voices or environmental sound sources in the auditory display, as the focus of active processing, reducing, or eliminating the competition.

The scientific interest in selective attention and its benefits for perception date to the late 1800s, when William James observed that an object that captured attention among many simultaneously possible objects seemed perceptually vivid and important. An introspective analysis of Wilhelm Wundt at about the same time asked the reader to fixate the eye on the center of an array of letters, while placing the subjective-fixation of attention on a letter some distance away. He noted that the attended letter and its immediate surround were more clearly perceived and letters away from the focus of attention seemed unclear and outside of immediate consciousness. These authors identified attention as an important, indeed core, aspect of perception. Their introspective observations presaged the concepts of *covert attention,* the dissociation of the focus of attention and the fixation of the eye, of *spatial attention,* a focus of attention on a region of space, and of *object attention,* the selection of focus based on objects. Yet almost a century went by between the introspective identification of these aspects of human perception and an objective, scientific understanding of the causes and consequences of selective attention in human information processing and the brain.

How does selective attention operate? Under what conditions is it important? How does attention affect the efficiency of information processing? Answering these questions about attention depends on a combination of clever selection of test conditions and an analysis of behavioral measures using formal or computational models of attention. Just as a physicist may arrange conditions to reveal a physical process, experimenters arrange test conditions to reveal mental processes. And just as mathematical theories assist in understanding many physical systems, so too are formal and computational models necessary for interpreting the behavior of human systems. Objective behavioral measures, such as the time taken to find a target, or the probability of finding a target in a very brief display presented for a fraction of a second, have replaced introspection and intuition in the study of attention.

This entry describes sensory processes and cued selective attention; selective attention in visual search; selective attention and control; selection, flanker effects, and crowding; and the significance of selective attention.

Sensory Processes and Cued Selective Attention

Sensory inputs arrive by stimulating sensory systems in the eye (or ear, or skin) that respond to specific patterns of light (or sound, or touch). Different cells in brain areas that process the sensory inputs respond to different patterns, and specific locations or objects often stimulate neurons in different locations in the brain. When you look at the visual world, sensory responses to the many stimuli across the visual field occur simultaneously, or nearly simultaneously, in parallel. They constitute a "blooming, buzzing" collection of sensory inputs or representations. Selective attention picks out the relevant from the irrelevant, the focus from the background. Many theorists have argued that attention selects one thing (or perhaps a few things) at a time for processing. So, they felt, this implied the processing of visual inputs serially—one after the next in time in a series. Selective attention may, however, function differently in distinct circumstances and is often best understood by modeling the ways in which attention operates.

Perhaps the strongest effect of selective attention occurs when a precue actively directs attention to a particular location or feature in advance. Researchers distinguish between two types of precues. Exogenous cues are flashes or other direct indicators at the location to be attended, and endogenous cues are cues that are not at the attended location. An example of an endogenous cue might be an arrow located at the center of a display asking the subject to attend to a separate, corner region. It takes between a tenth and a quarter of a second to process a precue and focus selective attention to a particular location in space.

Laboratory tests have shown that selective attention may speed a response to the stimulus but, more importantly, may improve the accuracy of identifying or classifying a briefly presented target pattern. When attention increases the accuracy of pattern identification, we can say it has improved the quality of processing. Knowing when and where to focus selective attention can be especially important when the target is obscured by noise, camouflage, or dense arrays of distracting elements. Models and tests of selective attention have shown two distinct ways in which attention may improve perception. One is to enhance responses to weak or low-contrast stimuli, rather like increasing the contrast on a television set to improve visibility. The other is to filter out other irrelevant items in noisy or cluttered displays and so improve the focus on the target. In both cases, the quality of processing is noticeably improved by directing the mind's eye of covert selective attention.

Selective Attention in Visual Search

Precueing of a spatial region focuses selective attention in advance. Researchers have also been fascinated with understanding the role of selective attention in searching visual displays in which the target location is unknown and may appear anywhere. The act of finding visual targets in complex visual scenes is a capacity that represents a common real world activity. Efficient visual search for environmental dangers likely had evolutionary significance. The processes—by which the natural targets of a lost child, the ripe fruit, or the proverbial tiger in the foliage are found—can also be tested in the laboratory. Visual search displays in scientific studies are designed to vary from simple to complex and generally consider sets of spatially separated objects.

For visual search, the primary theoretical question has been whether deployment of attention is serial, moving the focus of attention to one location or object after another, or can be done in parallel. In the laboratory, this involves looking for a single red target among distractors. Some cases, such as looking for the red target among all green items, can be direct and simple regardless of the number of items in the display. In these cases, the likelihood of finding the target in a brief display, or the time to find it in a longer display, is almost independent of the number of items. Searches where the target is immediately obvious regardless of the number of irrelevant distractors have been called *pre-attentive* (i.e., identified without attention) by some. In such cases, all the locations could be scanned in parallel, or at the same time, for the presence of the target feature, redness.

In more difficult searches, the target is not unique. An example is looking for a red horizontal line among distractor items that vary in color (red and green) and in orientation (horizontal and vertical). But this time, there are many red items and many vertical items, so the trick is to find the conjunction of the two properties together. One major question has been whether such difficult attentive searches require serial processing. If searching for a target among more distractor items takes longer or is less accurate, does this imply a serial process of attending to a sequence of items in turn? Not necessarily. If the job is to state whether a target was present in a display, even simultaneous parallel searches could show more errors or longer search times with, say, 10 distractors than with 2. Distractor items may be mistaken for a target or a target may be missed. Increasing the number of distractors leads inevitably to more missed targets or more false alarms. If there are too many false alarms, it makes sense to tighten the criterion for deciding that something is a target—thus increasing the likelihood of missing it. Furthermore, even parallel searches that begin simultaneously finish at different times. If no target is present, finishing the last of all of 10 items will be slower than finishing the evaluation of 2, even if all items are processed in parallel. So, parallel searches can show somewhat lower accuracy and longer search times with more distractors, but serial search models of attention predict more dramatic slowing as displays become larger. Other ideas about search include a combination of parallel and serial processes. Search may be serial, but may be guided by pre-attentive parallel evaluations that point to the more likely or salient items first, or eliminate some classes of distractors. For example, now the red horizontal target is hidden among a small number of red vertical items and a large number of green horizontal items. Often, people can use the comparative salience of the fewer red items to guide their analysis. These models have made it possible to more accurately understand how attention operates in different conditions.

The evidence indicates that parallel processing is more pervasive, and the role of covert attention selection more restricted, than was originally believed. The ability to accurately find a target in brief displays is quite often consistent with parallel processing of multiple objects, with added distractors causing exactly the predicted increases in errors. Parallel processing seems to account for the results of many tests, including many difficult visual search conditions that on the surface seemed to be limited by serial selective attention. However, serial processes may have a role to play in some difficult situations—those where targets cannot be seen well in the periphery, or where direct scrutiny is required to tell targets and distractors apart. Often, such difficult visual searches engage explicit movements of the eye to new positions in the field in addition to any deployment of covert selective attention. Such movements of the eye and associated episodes of attention are intrinsically serial. However, moving the eye is quite different from moving the internal focus of covert attention.

Selective Attention and Control

Selective attention is influenced both by goal-directed top-down control and by bottom-up processing of the visual or auditory inputs. Goal-directed control of attention occurs when a particular spatial region or feature is cued as more likely to include a target or when we know the goal of a search, or its features. A bird watcher may look for a bluebird by looking for blue features of a particular size in the trees. The focus of attention reflects needs and priorities, here the goals of the bird watcher. Conversely, selective attention may be triggered by salient elements in the visual (or auditory) field. A salient item may be one that suddenly appears, or a sole moving object, or an item that is unique in color or some other feature. Selective attention may be drawn without our volition to such unique elements within the field. Even when focused on blue features of the bluebird, attention may be automatically refocused by the sight of a red cardinal. These two mechanisms for directing the action of selective attention work together. Goal-directed selective attention can alter the sensory processing of the selected inputs, and unique features in the field may draw or redistribute selective attention.

Selection, Flanker Effects, and Crowding

Selecting desired target(s) and filtering out other inputs are the central role of attention. Selection by location, or by source, is a fundamental aspect of attention in both audition and vision. However, the ability to focus solely on a specific location is often imperfect. Nearby objects may influence behavior, especially those that are similar in content to the target objects. For example, if you are asked to classify the middle of three shapes, similar flanking shapes may become confused with the target. This is especially true when the focus of attention is away from the fixation of the eye. This phenomenon can easily be seen if you look at a page and try to identify the middle of three letters some distance from the fixation in the periphery. The spacing between letters matters, as does the similarity of the flanking items. For example, identifying a middle letter as an E or F will be more difficult if the surrounding letters are also Es or Fs than if they are Os. Again, peripheral targets are hard to see precisely when covert selective attention is insufficient for accurate identification and where eye movements must be used to move the potential targets closer to fixation during visual search.

Significance of Selective Attention

Selective attention is an important aspect of how we process incoming sensory information from the world. Deficits in attention, often measured as deficits in selective attention, are disrupted or altered in a range of mental conditions, including attention deficit disorder, schizophrenia, and stress and anxiety disorders, and is an aspect of aging. Further research in selective attention may have implications for understanding the changes in perceptual and cognitive processing in these conditions.

Barbara Dosher and Zhong-Lin Lu

See also Attention: Covert; Attention: Divided; Attention: Object-Based; Attention: Selective; Change Detection; Psychophysics: Detection

Further Readings

Dosher, B. (1998). Models of visual search: Finding a face in the crowd. In D. Scarborough & S. Sternberg (Eds.), *Invitation to cognitive science: Methods, models, and conceptual issues* (Vol. 4, pp. 455–521). Cambridge: MIT Press.

Dosher, B., & Lu, Z.-L. (2000). Noise exclusion in spatial attention. *Psychological Science, 11,* 139–146.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48,* 269–297.

Logan, G. (2004). Cumulative progress in formal theories of attention. *Annual Review of Psychology, 55,* 207–234.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40,* 1227–1268.

Serences, J., & Yantis, S. (2006). Selective visual attention and perceptual coherence. *Trends in Cognitive Sciences, 10,* 38–45.

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5,* 1024–1037.

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *TRENDS in Cognitive Science, 7,* 70–76.

Wundt, W. (1924). *An introduction to psychology* (R. Pintner, Trans.). London: George Allen and Unwin. (Original work published 1912)

ATTENTION: SPATIAL

One of the fundamental properties of our environment is that it is composed of a multitude of sensory information. Given such richness of input, humans are faced with the problem of having limited capacity for processing information, on the one hand, and the need to analyze as much of the sensory input as possible, on the other hand. Attention is the cognitive mechanism that allows effective selection of relevant information (e.g., the letters that you are reading right now) and an inhibition of "at the moment" irrelevant information (e.g., the humming sound of your computer).

Effective selection of information from the surrounding environment depends on the ability to select information in the most efficient manner and to distribute attentional resources in a systematic fashion based on the properties of the input. One fundamental property that describes our environment is its spatial nature—everything in our environment can be described by a set of unique three-dimensional coordinates (i.e., no two objects ever occupy the same location in space). Given that space is a unique descriptor, it is reasonable that the human attentional selection has evolved to use space as its fundamental unit of selection. Spatial orienting is not the only mode of attentional selection, however. Other units of attentional orienting are considered as possible candidates for attentional selection (e.g., object-based, feature-based, and modality-based orienting) on the grounds that spatial selection alone cannot account for the efficiency of attentional selection. This entry describes behavioral characteristics of spatial orienting, biased competition, attentional modulation, and sources of spatial attention.

Behavioral Characteristics of Spatial Orienting

Early models of attentional selection suggest that attention is directed to spatial locations in a manner analogous to a graded spotlight (or a zoom lens), selectively enhancing the perception of items that happen to occupy those spatial locations. An intuitive way to imagine how such a spotlight of attention might operate is to envision an image that has been blurred. An attentional spotlight is then moved around this blurred image bringing into focus spatial locations falling within its circumference. Consequently, items within the spatial spotlight are processed faster and more accurately than are those that fall outside of it.

One source of support for the existence of space-based attentional selection are the findings demonstrating that when a specific spatial location is attended, target detection at that spatial location is faster compared with targets that appear elsewhere on the screen. In such studies, subjects are typically seated in front of the screen and are asked to fixate their eyes on the small plus sign positioned at the center of the screen. Targets, flashes of light, or alphanumeric characters are then presented in various positions to the left and to the right of the fixation point. Participants are asked to respond with a button press once the presence of the target stimulus is detected. Of main importance is that the target flashes are preceded by a spatial cue—usually a flash of light—that is somewhat predictive (probability of 0.6 or higher) of the target location. Such design yields two types of cues: (a) valid—those that appear in the same location as the target, and (b) invalid—those that appear in locations other than the location of the target. The main finding from such behavioral cueing paradigms is that subjects respond faster to targets appearing in the same location as the cue (valid trials) compared with reaction times to invalidly cued targets. These results are interpreted in terms of a spotlight theory of attentional selection, such that spatial locations are selected in a manner analogous to a spotlight (i.e., only stimuli illuminated by the spotlight are selected), and that there are both costs and benefits to sharing attentional resources between spatial locations.

Space-based attentional orienting is not unique to the visual system (although it has been studied most extensively within the visual modality). Actually, early studies on spatial attentional orienting were conducted within the auditory domain. In a typical auditory spatial attention experiment, subjects are presented with two auditory streams (one in each ear, or dichotic presentation) simultaneously and are asked to shadow (i.e., repeat what is being spoken) one of the streams thereby focusing spatial attention on the sensory input within the left or the right ear. The first interesting observation emerging from such studies is that participants are able to perform this task at all! In addition, it is observed that subjects know surprisingly little (or almost nothing) about the ignored stream except that the sound is present. Subjects fail to notice such striking changes as language switches (e.g., from English to German), speech streams playing backward, or changing the gender of the speaker. Interestingly, not everything in the ignored stream is missed, and some stimuli can capture auditory spatial attention. In 1959, Neville Moray presented his subjects with a similar dichotic shadowing task, but in some cases, the experimenters inserted subject's own name in the ignored auditory stream. It was observed that the presence of one's own name could be detected within the ignored stream. As a result of such auditory spatial capture, however, subjects could no longer shadow the stream that was supposed to be attended. The results from these experiments suggest that in audition, just as in vision, stimuli that fall outside spatial attention are suppressed and fail to enter consciousness.

Biased Competition

The term *spotlight of attention* is merely a metaphor for changes in neural processing that follow spatial attentional orienting. The spotlight of attention is said to enhance an early sensory representation of the selected stimulus. Such enhancement can be viewed as a result of biased competition among neural representations. The biased competition model of selective attention proposes that items presented in a multielement scene are not processed independently, but rather interact in a mutually suppressive way. Such suppressive interactions have been observed behaviorally (e.g., harder to see an item when it is embedded among other items) as well as with neurophysiological and neuroimaging techniques. This neural competition, however, can be resolved via a biasing signal that is either bottom-up (e.g., red item among blue items) or top-down (e.g., looking for an apple in a supermarket's fruit aisle). To the extent that the biasing signal is spatially imprecise, its benefits spread within a local region, falling off with distance from the epicenter of the attended region. This could account for the distance effects as measured by both speed and accuracy in attentional paradigms in which items that appear closer to the cued location tend to be processed more efficiently (though not as efficiently as the valid location) than do those positioned further away.

The framework of biased competition consists of two general components: the source that generates the spatial bias and the effects of that attentional modulation bias on early sensory representations. The following two sections describe the neural mechanisms of each of these components in turn.

Attentional Modulation

When an item is attentionally selected (via the source signal), its representation is biased (i.e., enhanced) compared with the representation of this same item when it is not attended and/or is presented among other items (thus resulting in an attentional effect).

Studies of the effects of visual attentional selection demonstrated behavioral facilitation and enhanced cortical responses to attended locations. Neurophysiological studies investigating the effects of spatial attentional bias demonstrated that when a stimulus is presented in a neuron's receptive field (RF) the response to that stimulus is increased when spatial attention is directed to it compared with when attention is either unfocused or diverted elsewhere. For example, in their seminal study, Jeffrey Moran and Robert Desimone first identified the classic receptive field of a V4 neuron and its corresponding preferred and ineffective (not response eliciting) stimuli for that neuron. Monkeys were then trained to attend to stimuli in a specific location within the visual field while ignoring stimuli in other locations. Both effective and ineffective stimuli were then presented within the classic RF of a V4 neuron to elicit competitive interactions in a multielement display. The authors found that the responses of the V4 neurons were strongly modulated by the locus of the monkey's attention. The firing rate to the preferred stimulus was only one third as great when the monkey did not attend to it (i.e., attending to the ineffective stimulus) compared with when the preferred stimulus was attended. What is particularity interesting about this study is that when attention was directed to one of two stimuli in the RF of a V4 cell, the effect of the unattended stimulus was attenuated, as if the RF had contracted to only include the attended stimulus. Effects of attentional orienting are observed even when only one stimulus is present in the display.

Several neuroimaging studies, mainly event-related potentials (ERP) and functional magnetic resonance imaging (fMRI), also observed early sensory enhancement following spatial attentional selection. Studies employing neuroimaging techniques take advantage of the fact that visual stimuli presented to the left of the center are processed within the right primary visual cortex, whereas stimuli presented to the right of the center are processed within the left primary visual cortex. When one presents stimuli to the left and to the right of the fixation point, similar to the cueing paradigm discussed earlier, it is observed that directing spatial attention to the left hemifield results in the increased stimulus evoked neural activity of the early visual areas in the right hemisphere (and the converse is true for attention directed to the right hemifield).

Sources

Complementing studies that demonstrate the effects of attentional modulation is research that aims to determine the source of the attentional signal itself. Evidence from multiple methodologies has been accumulating that implicates a network in the posterior parietal and frontal cortices as likely candidates for the control of spatial attention.

Some of the first studies that investigated the role of the parietal cortex in the control of spatial attention have come from the neuropsychological literature. Certain forms of brain damage produce perceptual deficits in attending to regions of space. The most common of such deficits are extinction and unilateral visual neglect, both of which result from unilateral brain damage primarily in and around the parietal cortex. In extinction, the patient is able to respond to and recognize events and objects when they are presented in isolation anywhere in the visual field. When two items are presented simultaneously, however, one in the contralesional field (opposite to the location of the lesion) and one in the ipsilesional field, the contralesional object often is not reported or responded to. In unilateral neglect, the deficit is more severe in that even a single object in the contralesional field tends to be ignored in the absence of competition from other objects. Unilateral neglect tends to be a more debilitating condition in everyday life than does extinction, and some have argued that extinction amounts to a mild form of neglect. As such, the deficit in spatial orienting that is exhibited by unilateral neglect patients has been taken as evidence that the parietal cortex is involved in control of spatial attention.

Recent neuroimaging findings propose that control of spatial attention is likely subserved by the frontoparietal network of regions in the human cortex. Voluntary deployments of spatial attention are associated with neural activity in regions of the dorsal parietal cortex (intraparietal sulcus [IPS]), superior parietal lobule (SPL), and frontal eye fields (FEF). Conversely, the temporoparietal junction (TPJ) and ventral frontal cortex (VFC) are recruited when spatial attention is captured by a stimulus presented in an unexpected location (e.g., one's own name inserted within the unattended auditory stream).

Sarah Shomstein

See also Attention: Effect of Breakdown; Attention: Effect on Perception; Attention: Object-Based; Attention: Physiological; Attention: Selective; Attention and Consciousness; Visual Search

Further Readings

Bisiach, E., & Vallar, G. (1988). Hemineglect in humans. In F. Boller & J. Graffman (Eds.), *Handbook of neuropsychology* (pp. 195–222). Amsterdam: Elsevier.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25,* 975–979.

Connor, C. E., Gallant, J. L., Preddie, D. C., & Van Essen, D. C. (1996). Responses in area V4 depend on the spatial relationship between stimulus and attention. *Journal of Neurophysiology, 75,* 1306–1308.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Neuroscience, 3,* 201–216.

Kastner, S., & Ungerleider, L. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience, 23,* 315–341.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229,* 782–784.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11,* 56–60.

Yantis, S. (2008). Neural basis of selective attention: Cortical sources and targets of attentional modulation. *Current Directions in Psychological Science, 17*(2), 86–90.

Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., & Courtney, S. M. (2002). Transient neural activity in human parietal cortex during spatial attention shifts. *Nature Neuroscience, 5,* 995–1002.

ATTENTION: THEORIES OF

Every moment of our lives, the world bombards us with a multiplicity of sights and sounds and other information taken in through the senses. These stimuli may be relevant or not to your current cognitive goal, or of such potential importance that, unexpected as the input may be, you must break out of your current goal-directed activities to understand the new input. A mother might be looking for her child in the playground, but may still respond to the sound of her cell phone. An air traffic controller may be busy tracking a large number of display inputs while actively interacting with only one or two. These and myriad other examples reveal that we are exquisitely tuned to the ebb and flow of new information. Attention is the mind's solution to the problem of sensory and processing overload. Whether this involves the selection of a particular piece of sensory information for additional processing or carrying out two tasks at once, attention systems select inputs and coordinate processing resources of the human brain. Scientists have furthered their understanding of attention through both empirical and physiological observations and the development of theoretical models for mechanisms of attention, as described in this entry.

Early and Late Selection

Inspired by analogies between humans and computers, the first of the modern theories of attention tried to explain how humans selected some sensory inputs for further processing while ignoring others. These theories also emphasized the extent to which sensory inputs could be processed without attention. *Early selection* theories claimed that the attention system acted like a filter that could select some inputs for additional processing based on visual physical characteristics such as location, color, or texture or based on auditory characteristics such as location, pitch, or the speaker's unique voice. Processing for identification or memory was claimed to occur only for attended-to inputs—unselected inputs were filtered out completely. At the other end of the continuum, *late selection* theories argued that considerable processing could be carried out on many inputs without limitations, and that the bottleneck in processing occurred just before the choice and execution of a response.

Initial tests of these theories were similar in essence to the "cocktail party" phenomenon, where a listener must pick out a single speaker in a room full of competing conversations. Experiments present different messages, sometimes by different speakers, in each ear ("dichotic" listening) using headphones. Listeners might be asked to immediately repeat ("shadow") a target speaker, as the message is heard in one ear, while ignoring messages in the other ear. However, attention filtering was ultimately shown to be by no means complete or absolute. Unattended messages, such as those presented in the other ear, were at least partially processed, especially salient or high-valence messages, as in hearing one's name in the unattended message. People notice if a word in an unattended message is the same or related to the attended message, so long as they occur close together in time. Consistent with this, physiological measures have shown that the effects of attention seem to begin at the first stages of sensory processing. Event-related potentials (ERPs), or electric activity of the brain in response to visual inputs, show stereotypical differences in the responses for attended compared with unattended stimuli. Attention can modify the brain's responses to stimuli, amplifying relevant inputs or attenuating irrelevant ones, even the first and fastest responses. For example, the response over the visual cortex to an attended visual stimulus ("pay attention to targets in the upper right quadrant") causes a larger response within the first 100 milliseconds (ms) in the P100 wave of the ERP, and this extends through early visual responses in cases where attention focuses on one location over longer blocks of many trials. Ignoring inputs does not, however, eradicate the sensory responses. In short, a modified form of the early selection model provides a good account: selection affects the earliest forms of sensory processing, but ignored inputs are attenuated rather than filtered out completely. The magnitude of the difference in response between attended and unattended inputs depends on exactly how attention is controlled. Attention has smaller effects on the earliest brain responses in cases where the focus of attention is changed to a new location by a cue on each trial. So, the impact of attention on perceptual processing can be more profound with a sustained period of attending to one location or feature than if selective attention is shifted moment to moment by changing cues.

Filtering and Amplification Mechanisms of Attention

Early selection theories claimed that attention filters some stimuli out and enhances others, but they failed to do more than sketch the mechanisms of filtering or the consequences for information processing accuracy. However, two related mechanisms of attention have been proposed and tested within computational models of a human observer's performance. In these models, stimuli are processed through perceptual templates tuned to respond to targets. For example, a template might be tuned for the letter A, or a particular shape or sound. The response of the template is noisy, and this and other noisy processes limit the accuracy of identification. Separate mechanisms exist for the enhancement or amplification of an attended stimulus and the filtering out of external noise or distractors. The *template model of attention* specifies how filtering and enhancement separately affect the accuracy of detecting a target or discriminating between targets within a framework of signal detection theory. For visual tasks such as identifying a letter or the orientation of a patch of pattern, tests have been developed that measure identification accuracies in the presence of different amounts of noise in the stimulus, something like looking at a poor television image that has a lot of visual "snow." Here, cueing attention to a location has one widely observed mechanism that filters out the noise—it improves performance most precisely when noise or distractors would otherwise damage performance. A second less frequently seen mechanism, enhancement, improves performance in an attended location even in the absence of competing inputs and is especially important for weak targets. Attention enhancement and attention filtering can occur separately or together. These mechanisms can also be directly related to models of neural processing in the visual cortex.

Signal Detection Theory Framework

Signal detection theory (SDT) explains how performance accuracy depends on the strength of the target ("signal") stimulus compared with nontargets in the context of the variability, or noise, in the signal strength over test trials. SDT assumes a noisy response to each stimulus, and determines how criteria or decision rules classify the stimuli to make a response; it separates the true discriminability of target and distractor stimuli from possible changes in criterion. The SDT framework was the basis of the template model (discussed earlier) and has been important to understanding attention limitations in searching for a target among other items. SDT models make predictions about the accuracy of performance when people look for a target among different numbers of distracting stimuli or locations under conditions where people cannot achieve perfect accuracy.

One interpretation of poor performance with more distractors was that people have a limited capacity for processing many items at once. For example, early behavioral tests of these ideas involved briefly presenting a dim light in one of several locations, and noted that the accuracy of reporting the correct location decreased as the numbers of locations increase. Early researchers took these results at face value to say that it was not possible to process more locations without reductions in performance because of capacity limitations in attention. Today, researchers know that such conclusions require first an analysis of the role of chance guessing (or criteria) and require signal detection computations. Even if each item is processed as well in the presence of distractors as alone, the natural variability in the sensory registration of each item causes more errors with more distractors.

In the signal detection framework for visual search, each location (or item) yields an estimate of sensory strength, or match to a target template (the "signal"). Each estimate, whether from a target or a distractor, is noisy because of processing inefficiencies or variation in the presented stimulus. Each added input to the decision increases the potential for false alarms, or classifying a nontarget as a target in error. Even for an ideal machine in which all evidence is processed perfectly and recalled without capacity limits, added inputs reduce accuracy. Only when the losses in accuracy exceed those predicted for an ideal observer are capacity limits implicated. For almost all displays where a target is defined by a single feature or a simple combination of features, the detection or discrimination accuracy is consistent with the predictions of the unlimited-capacity ideal observer model. This includes simple detection ("did something appear?"), but also discrimination of features such as length, orientation, color, or simple shape ("was there a longer one?"). It also includes many difficult cases of visual search where the target stimulus is defined by conjoining two simple features, such as a given orientation and color, where the distracting nontarget stimuli may share either orientation or color with the target. In all these cases, the increases in errors with the number of display elements have been almost exactly those expected from signal detection computations of increasing false alarms (or reduced hits caused by increasing criteria to avoid those false alarms). Limited capacity seems to be limited to special cases where discriminating a target involves spatial relationships, such as two touching bars in which white is on the left and black on the right, or vice versa. Related models of the speed of decision also support unlimited capacity parallel processes, at least for the accuracy-limited brief displays.

Salience and Guided Search

Guided search or salience models of attention predict visual search time when it is possible to scrutinize displays with eye movements. Models of salience, which mark certain items for priority processing, have illuminated how attention may be guided bottom up from stimulus characteristics and top down through goals. Strong or unique sensory stimuli such as the only vertical item among all horizontal items, or one that appears suddenly, have high salience that makes a search easy if the salient item is the target, or make it more difficult if it is not. Stimulus attributes also can interact with behavioral goals to increase the relevance of some items over others, so if one is searching for a red vertical line, items sharing these features may inherit higher salience. In these models, salience of each item is computed in terms of bottom-up factors such as how different it is from other items, especially items that are close in the visual field. The salience is also determined by top-down factors such as the similarity of each item to attributes of a goal or target. For example, in searching for that red vertical line among red horizontal and green vertical items, the ease of search often depends, all else equal, on whether there is a small subset of red items (or vertical items) to which search can be guided.

Several models of salience-guided search have been developed and tested that provide algorithms for the computation of salience and predictions about how these salience computations interact with distinct behavioral goals and tasks. These salience models, and other forms of guided search model, have modified and improved a set of previous models that were based on the notion that the role of limited-capacity, serial attention processes was to bind back together, or integrate, the features of objects that were analyzed and taken apart by separate feature maps in early visual analysis. The original feature integration models of visual search made a contrast between visual searches without needing attention ("pre-attentive") and those that are sufficiently demanding that they need serial scrutiny. Pre-attentive searches occur when any single feature such as any red item independent of orientation, size, and other features is sufficient to define a target as distinct from all other cases where a conjunction or combination of several features (i.e., red and vertical) is needed to define the target. The search must locate an object that contains both features. However, more refined guided search models account for all the early results of visual search used to support the feature integration theory, but also account for findings that reject the feature integration model, such as the ability to restrict search to the subset of red stimuli. Predicting the time to find a target in a visual search task has been the primary goal of guided search and salience models. The salience computations may also make predictions about the likely focus of eye fixations during visual search. Salience-based attention and search models have also been developed in computer science or machine-vision applications, where image-analysis algorithms or programs that do salience calculations are developed to mimic human-like capabilities (and limitations) in visual search tasks.

Capacity and Scheduling

What happens when people try to carry out multiple tasks at the same time? Successful multitasking could magnify human efficiency. However, as failures such as the dangerous use of cell phones while driving or the inability to benefit from listening to different lectures in each ear show us, there are often important limitations to how much can be simultaneously processed. Models of attention capacity and bottlenecks have been used to explain these limitations in simultaneously carrying out two tasks, each with a decision and response. One approach measures *attention operating characteristics,* graphing performance measures from two tasks against one another while varying the instructed division of effort, from focusing entirely on one or the other and different sharing ratios (e.g., 70 to 30%), to show the extent to which two tasks interfere with one another. For example, accuracy of Task A (e.g., visual contrast increment detection) is graphed against the accuracy of Task B (e.g., an auditory contrast increment detection). Observing no sharing losses, that is, where the two tasks can be performed as well together as separately, corresponds to the absence of mutual capacity limitations, or the "ideal point," where the accuracies of Tasks A and B performed together with equal emphasis ("attend equally" or 50 to 50%) are as high as when each is performed alone. The "switching-line" connects single-task performances (from single Task A performance and chance Task B performance to chance Task A performance and single Task B performance), where performance improvements in one task trade directly with performance losses in the other, reflects choosing one or the other task, but not both, to do on any given trial. Often, performance is worse than the ideal point but better than the switching line, corresponding with some degree of successful sharing. The attention-operating characteristic provides a powerful measure of dual-task compatibility.

In other kinds of tasks, people are asked to carry out one task first, followed as quickly as possible by the other. Presentation of the two stimuli is offset by different amounts of time to measure the *psychological refractory period.* If asked to classify a briefly flashed letter by speaking the name and to classify a tone as high or low by pressing a key, the response to the second-arriving stimulus may be delayed. To be refractory is to be resistant, or unresponsive to a stimulus. If two stimuli occur close together in time, then the response to the second is often delayed as though it were necessary to process the first before there is capacity available to process the second one, often by as much as a second or more. In general, it has been concluded that initial perceptual analysis often is carried out on multiple inputs simultaneously, and the need for limited-capacity attention occurs later, perhaps at the level of decision or response selection. The extent of capacity limitation or interference depends on the demands of the tasks. There are two distinct interpretations. One is that bottlenecks of limited attention are intrinsic to certain stages, such as decision, and the other is that introduction of delays is a strategic choice designed to eliminate conflict at the level of the peripheral systems, such as the eye or the hand, which cannot be in two different places at once, and not a central capacity limit at all. This second theoretical framework has been tested using computer programs that simulate specific tasks and predict the pattern of response times. Whether obligatory or strategic, the existence of temporal delays in responding to two input stimuli close together in time is a general finding demonstrating people's limitations in responding to two tasks.

Control of Attention and Attention Switching

Shifting attention from one location to another is an essential part of how we view the world. Covert attention, movement of the mind's eye, leads and targets the movements of the eye. Covert attention shifts are also used to take up information in different spatial regions without movement of the eye. Instructing people to make shifts of covert attention is one strategy for measuring how attention is shifted from one location to another. In one classic example, an arrow points at and leads attention to a possible target location. The costs of misdirection, called invalid cueing, can be measured in lower accuracy and longer response time compared with valid cueing. In another example, individuals may be asked to search for a target in a stream of letters appearing in one location, one after the next, and then to switch immediately to a second stream to detect the next item. The theoretical metaphor of covert attention operating like a spotlight led some researchers to suppose that moving the mental attention spotlight would entail activating spatial regions between the current location and the new location. Clever theories and experiments, however, have shown instead that items in intermediate locations do not benefit from the shift of attention. Instead, if attention is redirected from one location to another, it is as if one spotlight is dimmed while another spotlight focused at a second location is lighted. Excellent quantitative models account for attention switches as such episodes of information acquisition through attentive deployment that operate separately in space and time.

As in the case of visual search, many of these tasks combine a goal-directed, or top-down, specification of a location or a target with a stimulus-based, or bottom-up, analysis of the items in the field and their salience. Indeed, it seems that many visual searches involve naturally evolving cycles of attention deployment that lead to eye movements, and then to a new reassessment of the search field, and another eye movement. In visual search, in planned sequences of shifts of covert attention, or in unplanned or evolving sequences of movements of the eye, the control and shifting of covert attention plays an important role in how information is acquired from the inputs of the outside world.

Physiological Substrates of Attention

With the explosion of new knowledge about brain structure and function, new physiological theories can describe the impact of attention on individual cells in the sensory analysis systems and in networks of centers that act to engage attention or to disengage and move attention, and to maintain vigilance. Spatial attention changes the strength, or rate, of response for individual neurons in the early visual cortex, especially in the presence of distracting stimuli. Individual neurons may respond as though distracting stimuli have been excluded from influence on the response, a property similar to signal-detection-based template models of attention as a filter. Individual neurons can increase the level of response to an attended stimulus as though the contrast or strength of the stimulus was increased. This property, sometimes called contrast-gain, is directly related to the concept of enhancement or amplification of the attended stimuli in quantitative signal-detection models of attention. Computational model frameworks that integrate these parallel models from the cellular level in one or another individual region of the sensory cortex up through full analysis and behavior remain to be developed. Still, it is clear that the behavioral and the neural responses share close functional analogies.

Distinct networks of brain centers have been implicated in goal-directed preparation for expected inputs or targets and in the reorienting of attention in response to salient or unexpected sensory stimuli. Frontal and parietal regions of the cortex, including frontal eye field and lateral intraparietal areas, are activated during goal-directed preparation or deployment of attention in response to a cue to a relevant location or feature. These systems are activated during the period when the expectations from a cue are maintained and may be recruited during goal-directed tasks such as visual search. A different network, one that is lateralized in the right ventral frontoparietal area, appears to be involved in the resetting and reorientation of attention when an unexpected or salient input overrides goal-directed orienting. In goal-directed situations, recent evidence suggests that frontal activity precedes activity in the relevant sensory areas. If attention is attracted by the stimulus, however, then activity in the sensory representations seems to lead to the frontal activity. These networks have been identified through new imaging technologies that sense the activated regions of the brain during attention tasks in humans, through behavioral anomalies in humans with lesions, and by measuring neural responses in tasks in monkeys that parallel those in humans.

Future of Attention Theories

Attention is a critical function in the mental arsenal, one that has fundamental implications for how we interact with the world. As we learn more about how the brain works, computational models could help link cognitive function to physiological responding, and psychophysical models of function and behavior to the workings of physiological components of the relevant brain networks or systems underlying control of attention. These new theories and models, in turn, would help us better clarify the role of attention. A better understanding of attention's "exquisite control" in determining which inputs are processed and when would further our understanding of intelligent information-seeking behavior and how we choose to manage multiple task demands. In turn, the understanding of the distinct mechanisms and brain circuitry of attention may also increasingly serve as a basis for understanding how the attention systems fail to function or function differently across the life span or in different mental health conditions such as attention deficit disorder or schizophrenia.

Barbara Dosher and Zhong-Lin Lu

See also Attention: Covert; Attention: Divided; Attention: Object-Based; Attention: Selective

Further Readings

Brefcynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the "spotlight" of visual attention. *Nature Neuroscience, 2,* 370–374.

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science, 315,* 1860–1862.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3,* 201–215.

Dosher, B. A. (1998). Models of visual search: Finding a face in the crowd. In S. Sternberg & D. Scarborough (Eds.), *Invitation to cognitive science: Vol. 4. Methods, models, and conceptual issues* (pp. 455–521). Cambridge: MIT Press.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48,* 269–297.

Logan, G. D. (2004). Cumulative progress in formal theories of attention. *Annual Review of Psychology, 55,* 207–234.

Lu, Z.-L., & Dosher, B. A. (1998). External noise distinguishes mechanisms of attention. *Vision Research, 38,* 1183–1198.

Pashler, H. (1998). *The psychology of attention.* Cambridge: MIT Press.

Shih, S-I., & Sperling, G. (2002). Measuring and modeling the trajectory of visual spatial attention. *Psychological Review, 109,* 260–305.

Treue, S. (2001). Neural correlates of attention in primate visual cortex. *Trends in Neurosciences, 24,* 295–300.

Wolfe, J. M. (2007). Guided search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York: Oxford University Press.

ATTENTION AND CONSCIOUSNESS

Have you ever arrived at work or a store and realized that you have no memory of seeing, hearing, or otherwise experiencing anything along the way? This is a sufficiently common experience (or lack of experience) that it has received a nickname in the perception literature. It is referred to as having been in the *zombie state.* The idea is that in this state, people behave as functioning beings with the capacity to take in and respond appropriately to sights, sounds, and other stimulation but do so while being unconscious of them. Unlike horror-movie images of walking corpses, this zombie state is not necessarily bad. In athletics, it is referred to as being *in the zone,* and it is often a goal state. For athletes and non-athletes alike, this apparently unconscious yet functioning state is most likely to occur when executing a skilled behavior such as driving a highly familiar route.

The idea of attention is often invoked in thinking about what controls whether one enters the zombie state. For example, if nothing unusual happens on your drive to work and you are thinking intently about something, such as a project at work or a problem at home, you are likely to be unaware of your surroundings. If, however, something draws your attention, such as a pedestrian darting across the road, you are likely to become quickly aware of your surroundings. Reputedly, one can knock an athlete out of "the zone" by getting him or her to attend to details of an action—"Do you grip more tightly with your left or right hand when making a long putt?" Both intuitively and empirically, there seems to be a strong link between attention and consciousness.

The link between attention and consciousness remains a controversial issue. Some theorists assert that attention is both necessary and sufficient for consciousness. The idea is that you must attend to something to become conscious of it, and if you do attend to it, then you will become conscious of it. This is tantamount to arguing that attention is the same thing as consciousness, in which case, for simplicity, it might be best to jettison one or the other construct. Other theorists assert that attention is necessary but not sufficient for consciousness. The idea here is that you must attend to something to become conscious of it but that attending to it will not always make you conscious of it. This position assumes that attention serves as a gatekeeper by which the cognitive system is protected from being overwhelmed by the vast amount of information that impinges on the sensory systems. Finally, still other theorists assert that attention is neither necessary nor sufficient for consciousness. Under this view, attention and consciousness serve distinct functions, but they interact with each other. What evidence is there for these different views? This entry will discuss some of this evidence.

Link Between Attention and Consciousness

Some of the earliest evidence for a link between attention and consciousness came from attentional cueing experiments pioneered by Michael Posner and colleagues. In these experiments, attention is directed to some location and awareness of information at that location is compared with awareness of information at other locations. A cue such as a brief flash of light might be used to direct attention to a location. Stimuli are then presented at that cued location or at uncued locations. Observers are relatively poor at identifying or otherwise indicating that they were aware of stimuli at uncued locations compared with cued locations, suggesting that awareness depended on attention.

Other evidence of a strong link between attention and consciousness comes from "selective looking" experiments. Here, two scenes are superimposed and observers are asked to attend to one while ignoring the other. For example, Ulric Neisser and Robert Becklen superimposed two movies—one of a pair of people engaged in a hand-clapping game such as "See My Playmate" and another of a group of people tossing a ball from one person to the other. When observers were asked to report on some aspect of one of the movies, such as how many times hands made contact or how many times the ball was passed from person to person, they were strikingly unaware of events in the other movie. In one case, for example, a woman holding an open umbrella walks through the group of people playing ball. Many observers who were monitoring the hand game failed to notice this odd event.

Exploring the link between attention and conscious perception further, Arian Mack and Irvin Rock designed experiments to ask whether stimuli are perceived at all when they are "truly" unattended, that is when attention has not been directed to them and when observers have no reason to suspect that they will be relevant to the task at hand. These experiments involved a difficult attention-demanding task. For example, observers would be shown a large plus sign in which either the horizontal or vertical bar was slightly longer, and the task was to report which was longer. Unbeknownst to the observer, after completing several trials of this difficult task, an additional stimulus was presented along with the plus sign, such as a colored shape. When asked after the trial whezther they noticed any stimuli other than the plus sign, observers often reported seeing nothing. This happened even when the additional stimulus was presented directly at the location where their eyes were fixated. Apparently, the plus-sign task prevented observers from attending to anything else and, as a consequence, concluded Mack and Rock, nothing else was perceived, a phenomenon they labeled *inattentional blindness*.

A related phenomenon, known as *change blindness,* seems to reinforce the link between attention and conscious perception. In these studies, which were pioneered by Ronald Rensink, Daniel Simons, and others, observers failed to notice large changes in a scene that they were viewing unless attention was somehow directed to the change. These experiments are similar to "find the differences" puzzles in children's magazines, except that the scenes are presented at different times instead of in different locations. For example, two versions of a city scene—one in which a (large!) steeple on a building is present and one in which the steeple is missing—are presented one at a time with a brief blank screen in between. Observers must find and report the difference across the two versions. This is a remarkably difficult task. In some cases, the scenes often cycle for more than one hundred times before observers reliably detect the change. Moreover, when monitoring eye movements during this task, it has been found that observers sometimes fixate the change without noticing it. Apparently, until attention is paid directly to the change itself, the experience is of an unchanging scene across the two views.

Finally, individuals who have suffered certain types of brain damage have also provided insight into the link between attention and consciousness. Patients who have suffered damage to parietal regions of their cerebral cortex, for example, which have been implicated in the control of attention, seem to show something like inattentional blindness, referred to as *neglect.* When attention is drawn to the areas of their visual field that correspond to the damaged brain areas, they are able to perceive stimuli normally. When attention is not drawn to these areas, however, the patients behave as though those stimuli are not there at all. A related phenomenon known as *extinction,* which also occurs with parietal damage, similarly highlights the link between conscious perception and attention. When a single stimulus is presented in a location that corresponds to the area of cortical damage, these patients do notice and can report on this stimulus. However, when two stimuli are presented—one in a location that corresponds to undamaged brain areas and one that corresponds to damaged areas—these patients will often report seeing only the stimulus in the "good" area, as if perceiving that stimulus "extinguished" conscious awareness of the other stimulus.

An important aspect of all these phenomena—attentional cueing, inattentional blindness, change blindness, neglect, and extinction—is that when attention is directed to the stimuli in question, observers are able to see them easily. Thus, it is not the quality of the stimuli that determines awareness, but rather whether or not they are attended. The question remains, however, is attention *both* necessary and sufficient for consciousness?

The Issue of Sufficiency

Several theorists have noted that although attention may be necessary for consciousness, sufficiency is yet a further step. Daniel Levin and Daniel Simons, for example, noted that observers are often unaware of changes that occur to aspects of a scene to which they are attending. Observers were asked to watch films in which various changes occurred, in anticipation of answering questions about the film afterward. Strikingly, about 50% of the observers failed to notice that the actor playing the main character in the film changed across a scene cut. In an even more dramatic example, observers came into the lab and met an experimenter behind a counter. After engaging the observer in a brief conversation, the experimenter ducked below the counter as if to get some materials on a shelf. Unbeknownst to the observer, a second experimenter who had been waiting below the counter, came up and resumed the conversation with the observer. Fewer than 40% of the observers in this experiment noticed this change.

Theorists who espouse the view that attention is necessary but not sufficient for consciousness tend to attribute a gatekeeping role to attention. The idea is that much more sensory information is available at any given time than can be fully processed. Attention controls the flow of information to later processes that are especially limited in their capacity. Consciousness is thought to arrive through access to some subset of these later processes. Thus, attention is necessary to achieve access to those processes, but it does not follow that all information that is let through will gain access to them. Stanislas Dehaene and Lionel Naccache, for example, codified this view into what they refer to as a *global workspace* framework following theorist Bernard Baars. The idea is that attention activates within a central representation—the workspace—information that is most relevant to the current contextual demands. Observers become conscious of some of this information when it engages some of the processes within this workspace, such as the establishment of a durable and explicit representation, or the initiation of intentional spontaneous behavior, but do not necessarily become aware of all of activated information. In this way, attention controls what one can be conscious of, but it does not directly determine conscious awareness of any information.

Perhaps Attention Is Neither Necessary Nor Sufficient for Consciousness

Although there is a strong link between attention and consciousness, some theorists have pointed out that this link may be neither necessary nor sufficient; attention and consciousness may simply interact with each other. Formally, for attention to be neither necessary nor sufficient for consciousness, one must occur without the other and vice versa. In two separate analyses of the literature, Christof Koch and Victor Lamme argued that exactly this situation holds. Examples were cited of every condition—attended and conscious, unattended and unconscious, attended and unconscious, and unattended and conscious. The first three categories constitute those reviewed in the preceding two sections of this entry. The fourth condition—unattended yet conscious—has been the most controversial and the most difficult on which to reach agreement: Can one be conscious of unattended information? The difficulty in answering this question is that assessing whether someone is conscious of something would seem to require asking about it, but in asking about it, attention is drawn to it, thereby destroying the necessary conditions for assessing whether consciousness is possible without attention.

Two strategies have been taken to circumvent the challenge of assessing consciousness without directing attention. One has been to seek indirect evidence that something has been perceived. Cathleen Moore, for example, presented the inducing portions of simple visual illusions in the backgrounds of displays using Mack and Rock's inattention method, which was described earlier. Despite being unable to report the patterns in the background, many observers nonetheless experienced the illusions that they were designed to induce. This suggests that those patterns were perceived at least sufficiently well to alter other aspects of the perception. Jeremy Wolfe coined the term *inattentional amnesia* to characterize this sort of experience and to contrast it with the term *inattentional blindness*. The idea is that observers may be immediately aware of sensory stimuli at the time they are present, but if attention is not drawn to it, it is not committed to memory, and the observer cannot report on it after the fact. The term *visual stuff* has been used by Wolfe, philosopher Wayne Wright, and others to characterize the type of perceptual experience that one has of unattended sensory information.

The second strategy for assessing perception without directing attention has been to use dual tasks. The idea is to engage observers' attention with a difficult task, and then ask them, secondarily, to report on other stimuli that might occur. Leila Reddy and colleagues used this strategy to ask whether people can identify faces without attention. Observers were asked to focus on performing a difficult letter-discrimination and search task in the center of the screen, while face stimuli were presented in peripheral locations. Reddy and colleagues found that observers were able to reliably detect the gender of the face with essentially no decrement in performance on the letter task and could identify individual faces with relatively little effect on the central task.

Conclusions from dual-task experiments always face the challenge that because observers were aware of the relevance of the other stimuli, they may have—however briefly—redirected attention to those secondary stimuli. Therefore, any awareness they may have of those stimuli could be attributable to residual attention. Because of this limitation, conclusions from dual-task experiments are usually limited to being of the form, for example, "gender identification in the near absence of attention," rather than "in the absence of attention."

Return of the Zombie

The difficulty of explicitly characterizing what is meant by "visual stuff" and the difficulty of ascertaining that no attention is paid to secondary stimuli in divided attention tasks constitute serious challenges to resolving the final critical question—Can people be conscious of things that they do not attend to? The view of theorists such as Koch and others, however, is that the zombie state reflects exactly this. It seems implausible that we drive to work, for example, with no perceptual awareness. Consciousness and attention are clearly closely linked. The final answer about the nature of that link will hinge on confirmation of the existence of the zombie state and an understanding of information processing while in that state.

Cathleen M. Moore

See also Attention: Covert; Attention: Disorders; Change Detection; Consciousness; Unconscious Processes

Further Readings

Dehaene, S., & L. Naccache (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79,* 1–37.

Koch, C., & Tsuchiya, N. (2006). Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Science, 11,* 16–22.

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Science, 7,* 12–18.

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin and Review, 4*(4), 501–506.

Mack, A., & Rock, I. (1998). *Inattentional blindness.* Cambridge: MIT Press.

Moore, C. M., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 339–352.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology, 7,* 480–494.

Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences USA, 91,* 7398–7403.

Reddy, L., Reddy, L., & Koch, C. (2006). Face identification in the near-absence of focal attention. *Vision Research, 46,* 2336–2343.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Science, 9,* 16–20.

Wright, W. (2006). Visual stuff and active vision. *Philosophical Psychology, 19,* 129–149.

ATTENTION AND EMOTION

Daily life presents such a bombardment of information that people would be overloaded without some means of prioritizing what they process. Attention and emotion systems both contribute to such prioritization. Emotions, for example, provide rapid, efficient means for identifying high-priority aspects of the environment, and attention mechanisms allow people to select manageable subsets from an otherwise overwhelming influx of information. Although these two systems influence each other, studies of attention within the traditional perception literature have often overlooked the role of emotion, examining instead how attention operates on various perceptual features. But the world is not characterized solely by assemblages of colors, angles, and motions; the objects, people, and events around us resonate with emotional meaning, so it is crucial to understand how attention and emotion interact. This entry describes emotional stimuli and varieties of attention, preattentive biases, rapid orienting versus delayed disengagement, emotional stimuli and mechanisms supporting awareness, asymmetry of attention-emotion interactions, and reciprocal influences.

Emotional Stimuli and Varieties of Attention

Attention refers to a family of mechanisms that—although they converge in the service of stimulus selection—differ from each other in important ways; orienting of attention to spatial locations is not the same as selectively attending to some features of a stimulus while ignoring other features, and neither of these processes is necessarily identical with the attention mechanisms involved in bringing information to awareness. Although evidence does suggest dissociations between these types of attention, each of them appears to be strongly influenced by emotion; emotional information seems to "capture" and hold various aspects of attention more robustly than does nonemotional information. For example, when it comes to attending to some features of a stimulus over others, *emotional Stroop* experiments have shown how difficult it is for people to ignore emotional aspects of a stimulus even when such aspects are task-irrelevant. In a typical version of this task, participants try to name as quickly as they can the colors in which words or monochrome pictures appear (or are printed). Frequently, they are slower to do so when the words and pictures happen to have strong emotional significance, suggesting that people had difficulty tuning out the task-irrelevant emotional information to focus only on the relevant color information.

A large portion of research on attention-emotion interactions has focused on the orienting of spatial attention. One procedure commonly used to tap into spatial orienting is the *dot-probe* task, where pairs of words or faces are typically presented on a computer screen and are followed quickly by a dot at one of the word/face locations;

participants are required to respond as soon as they detect the dot, and they tend to be faster when it appears at the former location of an emotional word or face than of a neutral one, suggesting that attention had already oriented to the emotional stimulus at that location (similar effects have been found even when people were not aware of the emotional stimulus). Similarly, in a *cueing* task—another measure of spatial orienting of attention—participants make speeded responses to targets, which could appear at one of at least two locations. On some trials, a cue appears before the target at one of the potential target locations, but the location of the cue does not predict the actual location of the subsequent target. In standard, non-emotional versions of this task, people tend to be slower to respond to the target when it appears away from the cue (an "invalid" cue) than when it appears at the same location as the cue (a "valid" cue), indicating that they had reflexively oriented to the cue despite knowing that doing so would not aid their performance. In emotional versions of this task, the cues themselves can be emotional or neutral stimuli (e.g., words or faces), and when they are emotional, their effects on spatial orienting are amplified. Notably, the emotional Stroop, the dot-probe, and various cueing experiments have revealed general biases to attend to emotional stimuli and have shown that such biases tend to be stronger among clinical and highly anxious individuals.

Pre-Attentive Biases

A number of models straddling the divide between cognitive and clinical branches of psychology have suggested that biases to prioritize emotional stimuli originate *pre-attentively,* meaning that they occur before attentional selection and are not necessarily bound by capacity-limited constraints. According to such models, pre-attentive evaluation of a stimulus's emotional significance helps direct the subsequent allocation of attention. Indeed, theorists such as J. Mark Williams, Fraser Watts, Colin MacLeod, and Andrew Mathews have suggested that high trait anxiety might be linked with a tendency to orient toward stimuli that were pre-attentively evaluated as threatening, whereas low trait anxiety might be linked with a tendency to orient away from such stimuli.

Some neurobiological evidence suggests how such a pre-attentive evaluation system might be instantiated, although such claims have also been challenged. For example, work on the amygdala—a subcortical structure strongly linked with the processing of emotional significance—has revealed direct connections with the visual system, suggesting neurobiological pathways through which emotional information could conceivably bypass many attentional circuits. Converging neuroimaging evidence has found heightened amygdala activity in response to emotional stimuli even when the stimuli were rendered unreportable through backward masking (i.e., when noisy visual patterns appearing immediately after emotional stimuli disrupted subjective awareness of such stimuli) and when attention was directed away from them to perform a secondary task. However, although such findings are provocative and have been widely influential, other work has found scant evidence of such heightened amygdala activity when the effectiveness of masking was rigorously ensured and when attention was strongly occupied by a secondary task. Additional studies have found that personality variables such as trait anxiety are linked with the degree to which attention manipulations modulate amygdala response to emotional stimuli. Such relationships may have been observed either because anxiety is linked with hyper-responsiveness to emotional stimuli or because anxiety is linked with a reduced ability to direct the focus of attention. Evidence exists to support both accounts.

Recent behavioral evidence also has led some to question whether emotional or otherwise high-priority stimuli are indeed processed independently of attention. In a study by Christine Harris and Harold Pashler, participants made speeded judgments about the relationship of two digits to each other and either an emotional word, the participant's own name, or a neutral word could appear in between the two digits. When the emotional word or one's own name was the only text alongside the digits, response times were slowed relative to when a neutral word appeared; this effect was especially large for one's own name. However, when one's name was only one of several words appearing alongside the digits, response times were no different from when no name was present. This finding suggests that rather than reflecting pre-attentive mechanisms, high-priority stimuli may receive preferential processing only when attentional resources are readily available. Of course, as the authors noted, it is possible that different results would have emerged had stimuli been even more emotionally evocative or had participants represented a more highly anxious population.

Rapid Orienting Versus Delayed Disengagement

Related to the question of whether emotional information can be extracted pre-attentively is the question of whether biases to attend to emotional stimuli reflect faster attentional orienting to emotional stimuli in the first place (which would be consistent with pre-attentive evaluation mechanisms) or delayed disengagement from emotional stimuli already at the focus of attention. Some evidence suggests the former. For example, one study used a *visual search* task in which participants searched for a fear-relevant stimulus (a spider or snake) among non-fear-relevant stimuli (mushrooms and flowers). Typically, in a visual search task, the time it takes to detect a target within an array of items increases as the number of array items increases. However, when the target was fear-relevant, response time seemed relatively unaffected by the number of array items (an effect known as visual "pop out"), suggesting that it had been among the first of the array items to draw attention. On the other hand, attentional cueing research in which cues could be neutral or threatening stimuli has suggested that rather than preferentially drawing an initial orienting response, emotional stimuli simply "hold" processing resources once they are attended. In one cueing experiment, when cues appeared at the location of the subsequent target, participants' response times were not affected by the emotionality of the cue, suggesting that the emotional cues did not elicit faster orienting. In contrast, when cues were invalid, appearing away from the target location, emotional cues led to slower response times, suggesting that participants indeed had difficulty disengaging from them to reorient attention to the target. This evidence suggests that when it comes to spatial orienting of attention, biases to attend to emotional stimuli may be driven by tendencies to linger on—rather than initially orient to—such stimuli.

Emotional Stimuli and Mechanisms Affecting Awareness

Recent studies have begun to examine the impact of emotional stimuli on attention mechanisms that help drive visual awareness itself. These experiments have suggested that emotional stimuli both gain more ready access to such mechanisms and, perhaps in doing so, prevent spatially or temporally neighboring, non-emotional information from doing the same. In large part, such experiments have used what is known as the *attentional blink* task, where participants search for targets within rapid streams of stimuli (e.g., streams of alphanumeric characters in which each item appears for about 80 to 100 milliseconds [ms] before being replaced by the next). In a typical non-emotional version of this task, people often detect the first target but fail to detect the second if it follows too soon afterward. One widely influential explanation for this effect is that the attentional processes that select information for consolidation into visual awareness are relatively slow, and that failures to detect the second target stem from such processes already being engaged by the first target. Notably, when the second target happens to be an emotional word, it is less susceptible to the attentional blink, suggesting that it captures the processes necessary to support awareness (although this does not occur among patients with bilateral amygdala damage). In a similar rapid presentation task, participants searched for only a single non-emotional target, and the target could be preceded in the stream by a task-irrelevant emotional or neutral picture. When the task-irrelevant picture was emotional, participants had difficulty perceiving the subsequent target; thus, in drawing attention to themselves, emotional stimuli appear to distract or disrupt attention processes that would otherwise usher other information into awareness. Consistent with this notion, evidence suggests that emotional stimuli associated with disrupted target detection are, in themselves, better remembered than are those not associated with such disruption.

Asymmetry of Attention–Emotion Interactions

The literature on attention–emotion interactions seems to contain an asymmetry wherein emotionally negative stimuli influence attention more than emotionally positive stimuli do. Such findings have led to suggestions of evolved mechanisms for attending specifically to threat; arguably, such mechanisms could aid one's chances of survival. However, some have argued that the observed asymmetry stems not from evolved threat-detection mechanisms, but from the fact that emotionally negative stimuli tend to be more emotionally "arousing" or intense than are emotionally positive stimuli. A few studies have attempted to control for the general emotional arousal elicited by stimuli and, in doing so, have reported that the positive–negative asymmetry largely disappears. In other words, it may be that the degree of arousal elicited by a stimulus drives attention effects more than does the evaluation of a stimulus as being positive or negative. The degree to which arousal and positive–negative evaluations respectively contribute to attention–emotion interactions is still a topic of considerable inquiry.

Reciprocal Influences

In addition to evidence suggesting that emotion influences attention, recent evidence has highlighted the bidirectional nature of this relationship, exploring ways in which attentional withdrawal from stimuli affects emotional responses to them. Jane Raymond and colleagues, for example, have found that when a visual search task requires that people ignore otherwise attractive visual distractors, participants later rate the ignored distractors as being less pleasing than visually similar items that had not been ignored. Notably, the harder a person had to try to ignore a distractor (e.g., the closer a distractor had appeared to a target), the more they affectively "devalued" it. Such findings reveal reciprocal influences between attention and emotion and carry implications for how patterns of attention allocation may foster affective preferences in daily life.

Steven B. Most

See also Attention: Covert; Attention: Effect on Perception; Attention and Consciousness; Emotional Influences on Perception; Face Perception; Individual Differences in Perception

Further Readings

Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature, 411,* 305–309.

Arnell, K. M., Killman, K. V., & Fijavz, D. (2007). Blinded by emotion: Target misses follow attention capture by arousing distractors in RSVP. *Emotion, 7,* 465–477.

Compton, R. J. (2003). The interface between emotion and attention: A review of evidence from psychology and neuroscience. *Behavioral and Cognitive Neuroscience Reviews, 2,* 115–129.

Fenske, M. J., & Raymond, J. E. (2006). Affective influences of selective attention. *Current Directions in Psychological Science, 15,* 312–316.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General, 130,* 681–700.

Harris, C. R., & Pashler, H. (2004). Attention and the processing of emotional words and names: Not so special after all. *Psychological Science, 15,* 171–178.

Mogg, K., & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour Research and Therapy, 36,* 809–848.

Most, S. B., Chun, M. M., Widders, D. M., & Zald, D. H. (2005). Attentional rubbernecking: Cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin & Review, 12,* 654–661.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General, 130,* 466–478.

Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences USA, 99,* 11458–11463.

Schimmack, U. (2005). Attentional interference effects of emotional pictures: Threat, negativity, or arousal? *Emotion, 5,* 55–66.

ATTENTION AND MEDICAL DIAGNOSIS

When a radiologist is presented with a medical image, be it a radiograph or the many hundreds of images generated from a computerized tomography (CT) or magnetic resonance imaging (MRI) scanner, he or she needs to make sense of the images, which are representations of the human body, and perceive pathology among the different ambiguous shapes, shades, and contours. Abnormalities are generally perceived quickly, as eye tracking has demonstrated, with pathology usually looked at within the first two or three fixations. It takes many cases and years of training to become proficient in interpreting medical images as perceptual discrimination is learned and acquired knowledge is converted into a variety of cognitive strategies and cognitive skills. This process is still not fully understood, but it seems practice with feedback is the only way to achieve clinically acceptable performance. Medical imaging has seen many developments in equipment, including the recent move away from film to digital technologies. Research has, however, demonstrated that error rates during the past 50 years remain resistant to these improvements and changes in technology. This demonstrates the importance of perceptual and cognitive factors in the performance of radiologists when interpreting medical images. This entry discusses eye movements, models of medical image perception, searching medical images, and decision processes.

Eye Movements

Eye tracking as a research methodology has provided some insight into the strategies and types of errors that are made by radiologists. Eye movements are not involuntary in visual search but can be described in terms of target selection, which in turn is related to the motivational state of the radiologist and to higher cognitive processes.

Errors can be search errors because of incomplete scanning, detection errors caused by failure to recognize visual features, and decision errors where the wrong decision is made. Most errors, however, are not caused by a failure to perceive but by a failure to recognize and make the correct decision. In mammography, about 30% of cancers are missed, and 70% of those missed cancers are actually looked at, although not identified as cancer.

Eye tracking has demonstrated different search patterns between novices and experts. For example, when looking at a chest radiograph, the expert will exclude large parts of the image during a search for lung nodules and concentrate on regions where lung cancer is more likely to occur. Compared with the novice, the expert will also make fewer fixations and have a greater distance between fixation clusters.

Reflecting the fact that there is currently no formal theory of optimal eye movement strategy in conducting visual search, there is no optimal strategy for radiologists, but eye tracking is useful for attempting to understand the way radiologists assimilate information from a medical image.

Models of Medical Image Perception

The global focal model of medical image interpretation, developed by Harold Kundel and Calvin Nodine, can actually be applied to any situation where domain-specific knowledge is important. According to this model, two forms of image analysis are performed sequentially. The first stage is the global or holistic impression (see color insert, Figure 12). This occurs instantaneously, in the same way as recognition of a familiar face happens. Any perturbations in the image are identified based on learned templates, and are flagged for subsequent searching before the first fixation. This is why abnormalities are usually looked at within the first few fixations. The second stage is when focal scanning takes place, which is essential for subtle lesions that are less conspicuous.

With experienced radiologists, the concept of normal is so highly developed that normal features are not consciously processed, but the abnormal ones are identified for focal scrutiny, so visual attention is perceptually attuned to the detection and decision making requirements of the task. For example, in mammography, where the appearances of some cancers can be subtle, the emphasis is on the cognitive aspect of visual search, which depends on training and expertise.

Searching Medical Images

There is a great deal of variability between radiologists in the way medical images are searched, but despite the variability, search patterns are rarely random. The relative importance of the image properties or cognitive structures in driving search depends on the task. Medical image interpretation is essentially task driven, where knowledge structures are more important than visual saliency. This is perhaps best illustrated by the way that clinical history or knowledge of the clinical indications for the imaging test can change the allocation of attention to different parts of the image, so perception can be changed in response to information about the implied meaning of the image. For example, knowing the patient's symptoms and the mechanism of injury can facilitate the detection of subtle fractures, such as some injuries to the pelvis commonly suffered by long jumpers, hurdlers, and gymnasts. Medical images are also representations of reality so perceptual organization depends on the observer's ability to retrieve stored mental representations of targets they may encounter.

One phenomenon that affects performance is the satisfaction of search effect (SOS) in which an abnormality is missed because another abnormality has been detected. Search is terminated once an expected pathology is found. This has been extensively researched by Kevin Berbaum in a recent study in which fractures were added to the images of 70 simulated multi-trauma patients leading to a substantial reduction in the detection of subtle fractures by radiologists.

Decision Processes

Once an abnormality is perceived, the decision about whether to report it is not always clear-cut. Radiologists will have different thresholds of abnormality depending on factors such as the task or even personality traits such as the cautiousness of the radiologist. For example, in screening examinations such as mammography where the occurrence of disease is rare (approx. 0.5%) and the consequences of missing a cancer are serious, then the radiologist when deciding on any abnormalities perceived in the image is likely to err on the side of caution and suggest a biopsy.

Computer-aided diagnosis (CAD) has been used to support the radiologist in his or her decision making by indicating any potentially abnormal locations that the radiologist might not have perceived and so reduce the number of missed cancers. Studies generally indicate that there is an improvement in detection with CAD, but not always, because not all observers use the CAD information in the same way, particularly if the CAD information conflicts with the physician's perceptual processes. For example, CAD can be used in mammography as a prompting system by highlighting suspicious areas for the observer to assess; however, cancers are rare in mammography so most prompts would need to be ignored, which, apart from being distracting, will also reduce the threshold a film reader will put on any prompt that it is a true positive.

Tim Donovan

See also Attention: Effect on Perception; Eye and Limb Tracking

Further Readings

Berbaum, K. S., El-Khoury, G. Y., Ohashi, K., Schartz, K. M., Caldwell, R. T., Madsen, M., & Franken, Jr., E. A. (2007). Satisfaction of search in multitrauma patients: Severity of detected fractures. *Academic Radiology, 14*(6), 711–722.

Krupinski, E. A. (2006). Technology and perception in the 21st-century reading room. *Journal of the American College of Radiology, 3,* 433–439.

Kundel, H. L. (2000). Visual search in medical images. In J. Beutel, H. Kundel, & R. L. Van Metter (Eds.), *Handbook of medical imaging: Vol. 1. Physics and psychophysics* (pp. 837–858). Bellingham, WA: SPIE Press.

Nodine, C. F., & Mello-Thoms, C. (2000). The nature of expertise in radiology. In J. Beutel, H. Kundel, & R. L. Van Metter (Eds.), *Handbook of medical imaging: Vol. 1. Physics and psychophysics* (pp. 859–894). Bellingham, WA: SPIE Press.

ATTENTION AND MEMORY

Behavior in a variety of visually guided tasks—from remembering a few objects to executing complex activities—is strongly influenced by the interplay between attention and memory. This entry highlights these interactions as we process our visual surroundings. Reflecting the corpus of scientific literature, the primary focus of this entry will be on short-term memory, although some interactions with long-term memory will also be described. Specifically, the focus is on the role of attention in determining the contents of visual memory, and, reciprocally, the role of visual memory in controlling attention during perceptually guided tasks.

As the following sections show, the likelihood that an object is represented in visual short-term memory (VSTM) is related to the amount of attention it receives during initial perceptual processing. Once in VSTM, attention-based rehearsal mechanisms affect the likelihood information is retained there or transferred into a more durable long-term memory (LTM) store. Retained short- and long-term memory representations of object location and identity can then reciprocally influence the subsequent guidance of attention through the visual world.

Encoding of Information Into Visual Short-Term Memory

As its name implies, visual short-term memory enables the temporary retention of newly acquired visual information over short periods (seconds). For example, VSTM allows one to keep track of car positions and movement trajectories between glances to the rear-view mirror while driving. In contrast to the seemingly limitless capacity of long-term memory, however, VSTM is a limited capacity resource restricted to, at best, four items—only a tiny percentage of the world around us (long-term memory enables an observer to accumulate information beyond this limit). With such a limited ability to store shortly relevant visual information in the short term, some set of mechanisms must govern access to VSTM. Although the importance of attention in determining the contents of VSTM has been known for some time, it is highlighted by two phenomena that have received recent consideration: change blindness and just-in-time search strategies.

Change blindness is the tendency for an observer to fail to notice changes introduced to a scene such as color alterations, object additions or deletions, and object substitutions from one view to the next. To induce this effect, one only needs to temporarily block an observer's view of the world when the change takes place. A lab-based example might be a momentary flicker on a computer screen and a real-world example might be a bus moving across a pedestrian's line of sight. These failures in visual memory and visual awareness result from a limited ability to efficiently encode the myriad objects and features contained in a scene into VSTM. Change blindness is reduced, however, when changes are made to objects that are likely to be attended. For example, changes to objects that are of central interest in a scene are detected faster than are those of more marginal interest. Indeed, once a change is found, it becomes so obvious to observers that they are surprised the task was so difficult.

Performance on perceptually guided action tasks such as making a sandwich, preparing tea, or sorting blocks, also suggests that attention is in some sense a gatekeeper for VSTM. When executing these various natural actions, observers adopt a *just-in-time strategy* in that they only obtain and represent information in VSTM when it is needed. For example, when asked to arrange a random assortment of colored blocks to match a model, observers typically fixate (attend to) a block in the model before picking up a block of the same color only to then look back at the model before placing the selected block in their copy. Observers apparently only acquire color in the first fixation and have to re-attend to the model to obtain location information in the second fixation. Hence, in perception-for-action situations, specific momentarily relevant features of an object or scene are selectively attended and retained in memory at the expense of task-irrelevant features. In conjunction with studies on change blindness, results of this kind suggest that focused attention may be necessary before an object (or a subset of its specific component features) is encoded into VSTM.

Maintenance of Information in Visual Short-Term Memory

Given its limited duration and capacity, the contents of VSTM must be constantly updated or "refreshed." Attention also plays a role (although not a solitary one) in determining the degree to which information is maintained in VSTM. Consider a situation where an observer is asked to remember a set of objects across a short retention interval. If at the same time, observers are also engaged in a difficult attention-demanding task such as counting backward from a random three-digit number in multiples of three, they experience a decrease in the amount of information that they can retain about the objects. Similarly, if a sudden visual distraction occurs during the retention interval, such as a localized flash of light, memory for objects previously presented near the distracting event are remembered better than are those that were further away as though a "spotlight beam" of attention shifted and narrowed from all objects to just a few. Finally, the number of objects that a particular individual can remember is directly related to his or her ability to vigilantly attend to task-relevant information and to ignore task-irrelevant information. Collectively, these results have led researchers to postulate attention-based rehearsal mechanisms that maintain information in VSTM.

Guidance of Attention

To interact with the visual environment, whether to find a misplaced book, to make lunch, or to safely maneuver an automobile, observers shift their attention from place to place. Although the guidance of attention through a visual display is likely influenced by perceptual factors such as local luminance, contrast, color, and motion, cognitive factors including memory are also important. For example, the contents of VSTM can bias the deployment of attention toward some objects and away from others and thereby influence the ease with which objects can be located and identified. To illustrate this, the following discussion focuses on the role VSTM plays in visual search tasks by considering how VSTM for spatial location and object identity affect attention guidance.

Memory for Object Location

Visual search is more efficient if an observer can remember the locations he or she previously searched and avoid revisiting them in the future. One method that has been used to explore the role that memory for object location plays in selecting items to interrogate has been to engage observers in a search task in which items are revealed a few at a time. For example, in the preview or *gap paradigm,* a search array is revealed in two stages. First, a set of distractor objects is presented, all of which the observer knows are not targets. After a delay of about 1 second, another set of items, one of which is the target, is added to the display. When objects are presented in this manner, the target is found faster than in situations where all search items are presented simultaneously. In order for this preview effect to be observed, some memory for the locations of the old and new items must be involved because once the second array appears, it is impossible to perceptually distinguish the two groups of items.

In addition to showing that attention is biased away from previously attended locations, researchers have also shown that visual search is slowed when observers are asked to remember a series of spatial locations during a visual search task. This occurs because VSTM is limited in capacity and the additional task interferes with the observers' ability to remember the locations of the searched and the to-be-searched items in the display. Additionally, these spatial memory loads increase the likelihood that attention will return to previously examined objects, thereby decreasing the efficiency of visual search.

Memory for Object Identity

When engaging in a visual search task, observers have to remember what they have looked at, as well as what they are looking for. Many models of visual search suppose that this is accomplished by generating a *target template,* or a representation of what one is looking for, in VSTM. Several lines of experimental evidence suggest that once this is done, attention is biased toward objects that are similar to this stored template, especially when the identity of the target changes from trial to trial (although this biasing is not automatic or mandatory, indicating that the contents of VSTM can be used flexibly across various task demands to guide the allocation of attention). In addition, after memorizing a given object, one's ability to locate and identify a search target in a visual array is facilitated if the target appears near an item that matches the to-be-remembered object compared with situations where the target appears near a visually distinct distractor. This memory-guided allocation of attention by object identity improves overall discrimination times and can attenuate set size effects (the tendency for search time to be directly related to the number of objects in the display) during serial visual search.

The bias of attention toward objects that match the contents of VSTM is not confined to serial search where observers must attend to each object in a display one-at-a-time. Even in a pop-out search task where observers are to report the shape of a color singleton target (e.g., the red target among green distractors or vice versa), search is faster if the target on trial N and N–1 are the same color compared with cases where their colors mismatch. By varying the lag between matching target trials, this priming of pop-out effect has been shown to persist for approximately 30 seconds.

Although explicit knowledge of the target identity facilitates search, recent evidence suggests that VSTM can help guide attention even when the identity of the target is unknown. For example, if one asks an observer to search for a suddenly appearing object in a real-world scene—and assuming that these objects are onset so that low-level motion signals that would otherwise capture attention are eliminated—the speed with which the targets are overtly attended increases with additional viewing time before the onset of the target. This result suggests that as more information is added to VSTM as viewing progresses, the resulting representation becomes more functional in the sense that it can be used to help orient attention within scenes even when precise search templates cannot be formed in memory.

Visual Long-Term Memory

In contrast to short-term memory, visual long-term memory has no clear limit on its capacity or storage duration. In one striking demonstration of this, after being shown hundreds of photographs for just a few seconds each, observers are able to recognize the pictures they saw even after a retention period of an entire year. Access to this LTM store relies on attention in much the same way as VSTM because the objects that are attended during the perception of the environment and that are rehearsed in VSTM may be retained long-term. However, the long-term maintenance of information in memory is accomplished without the need for sustained attention. As the previous example shows, even with long delays between study and test where attention has been removed from an object, its representation in LTM perseveres.

In addition to enabling the accumulation of information in memory beyond the limits of VSTM, long-term memory can also modify the allocation of attention to a visual display. For example, if an observer is repeatedly exposed to a constant arrangement of target and distractor items in a visual search task, search becomes progressively more efficient, an effect known as *contextual cueing*. This increase in efficiency occurs because repeated experience with a display leads to long-term memory for the configuration of objects that in turn can help guide visual attention to behaviorally relevant targets. That is, when the spatial relationships among objects in a scene are relatively constant, such as the arrangement of furniture in your living room, by virtue of knowing the locations of any set of objects, one also knows (or at least can predict) the location of any other item, thereby reducing the need to execute a detailed serial search for some desired target.

Long-term memory for object or scene identity is also used to guide attention during visual search. For example, target detection is facilitated if a target object is consistently presented with the same distractor objects, compared with situations where the distractor and target objects are uncorrelated. Additionally, the visual knowledge stored in LTM regarding semantic associations between objects influences search. When searching for a target object (e.g., nails) that is—unknown to the observer—absent from a display (hence the target cannot attract attention itself), the presence of a semantically consistent object (e.g., hammer) draws attention because, in our experience, these objects tend to be co-located. Thus, the visual information accumulated over one's lifetime can be brought to bear on search tasks in an effort to increase efficiency. Finally, returning to contextual cueing, if a learned configuration of objects is unexpectedly mirror-reversed so that they are spatially translated across the display, observers first shift their gaze (attention) toward the position in the display in which the target had previously appeared. However, when this initial search fails, attention is quickly directed to the target's new position, suggesting that LTM for scene and object layout are also used to guide attention.

James R. Brockmole

See also Action and Vision; Attention: Cognitive Influences; Attention: Effect on Perception; Change Detection; Eye Movements: Behavioral; Visual Memory; Visual Search

Further Readings

Awh, E., Vogel, E., & Oh, S.-H. (2006). Interactions between attention and working memory. *Neuroscience, 139,* 201–208.

Brockmole, J. R. (Ed.). (2009). *The visual world in memory*. Hove, UK: Psychology Press.

Chun, M. M. (2003). Scene perception and memory. *Psychology of Learning and Motivation, 42,* 79–108.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences, 9,* 188–193.

Simons D. J., & Rensink R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9,* 6–20.

Woodman, G. F., & Chun, M. M. (2006). The role of working memory and long-term memory in visual search. *Visual Cognition, 14,* 808–830.

ATTRACTIVENESS

The term *attraction* can refer to numerous interpersonal factors from liking to love, but it is most commonly used to indicate evaluative judgments of physical attractiveness. Although now prominently featured in journals of many disciplines, the scientific study of attraction is a relative newcomer to psychology. Since its beginnings, however, numerous perceptual cues have been identified that reliably affect perceptions of attractiveness. This entry reviews theoretical approaches to the study of attractiveness and then describes the visual, auditory, and olfactory cues that affect perceived attractiveness.

Theoretical Perspectives

Social Beginnings

The original and long-standing approach to the study of attractiveness was purely social. Classic studies examined the effects of physical proximity and interpersonal similarity to determine which factors affected interpersonal attractiveness and the consequences thereof. Two effects have emerged as important components of interpersonal attraction. First, the effects of propinquity, or closeness in physical proximity, on attraction are clear. People are more attracted to those with whom they have the least functional distance, and this is true for both friendships and romantic attractions. This occurs because of multiple factors including availability and mere exposure. Second, the effects of similarity are also clear, and they emerge along several dimensions. Overall, people tend to be attracted to others who share similar interests and personalities. Additionally, people tend to pair-bond with others who are similar in the degree of physical attractiveness (called *assortative mating*). Empirically speaking, attraction is more likely to follow the adage "birds of a feather flock together" instead of "opposites attract."

Other classic studies examined the consequences of perceived attractiveness. Attractive individuals, for example, enjoy what is called a *halo effect.* When someone is perceived to be attractive, observers are also likely to perceive the person to be likeable, successful, and socially skilled. Moreover, attractive individuals elicit more positive behaviors from others. These effects begin early in life. Infants prefer to look at more attractive others—including men, women, and even other infants! Despite the numerous benefits that beauty affords, it would be incorrect to assume that the effects of perceived attractiveness are entirely positive. Attractive individuals are also likely to be perceived as bourgeois, vain, and more likely to divorce, among other things.

All told, the early work investigating perceptions of attractiveness, though interesting, presumed that cues to attractiveness were socially constructed. Recently, however, this perspective has been largely supplanted by research that focuses on the biological, rather than social, determinants of perceived attractiveness.

Evolutionary Perspective

The interest in biological determinants of perceived attractiveness is typically described as evolutionary psychology. Researchers who adopt this perspective propose that human preferences reveal the distal determinants of mate selection. What is perceived to be attractive, it is argued, may also be adaptive. Thus, perceptions of attractiveness reveal mechanisms governing mate choice, and specific preferences increase the likelihood of getting one's genes into future generations. This perspective recently regained the imagination of researchers across disciplines, and research examining such questions proliferated.

Although the theoretical perspectives that generated the extant research base may vary, each finding in the literature rests on social perception in some respect. Indeed, numerous cues—including visual, acoustical, and olfactory—have been implicated for physical attraction. This entry now turns to these three distinct modalities of perception and highlights the relevant findings within each.

Physical Determinants of Attractiveness

Visual Cues

Most perceptual research focused on attractiveness has examined the visual cues that observers find attractive. Such cues are plentiful, and they appear from head to toe. In many cases, similar principles apply to the perception of bodies and faces.

Symmetry

One factor that is strongly linked to perceptions of attractiveness is bilateral symmetry in the face and body. Symmetry is argued to be an honest indicator of health. Because exposure to pathogens during development disrupts physical development, only those who experienced low pathogen exposure embody physical symmetry. Indeed, studies have linked bilateral symmetry to both health and fertility outcomes, suggesting that preferences for symmetric others may serve adaptive goals.

At times, symmetry can be perceived directly. Yet, bilateral symmetry need not be noticeably apparent to affect perceived attractiveness. Most research that has examined the effect of symmetry on social perception has measured more inconspicuous regions of the body (e.g., the length of earlobes or the circumference of wrists). Even when symmetry is less visually prominent as in these cases, it nevertheless affects attractiveness ratings reliably. Moreover, symmetry—and the health that it is theorized to signal—is also revealed in the perception and evaluation of other cues such as dancing ability (more symmetric men are better dancers) and body scent (symmetric people smell better). In such cases, it is unlikely that symmetry, per se, affects dancing ability or body aroma, but rather that healthy individuals tend to exhibit a set of cues including symmetry, physical abilities, and body odors, that co-occur and are perceived to be attractive by observers.

Neoteny (“Baby Face” Appearance)

Cues to youth are perceived to be attractive, and this is especially true of women. Many of these factors include infantile features and are apparent in the face. For example, faces with a youthful appearance include wide-set and large eyes, small noses, and smooth texture. These cues change dramatically with age, and they are perceived by observers to be attractive. Similarly, hair coloring darkens with age (before possibly going gray, which is a distinct cue to advancing age), and lighter hair is also judged to be attractive. Because perceived attractiveness of such features is more potent for judgments of women, scholars have argued that such preferences promote the selection of fecund mates.

Averageness

It at first seems odd to say that averageness is attractive. Yet a considerable amount of research has found just that, though perhaps not in the way that comes to mind immediately. “Average” images are not images of the most average members of society, but rather composites of several faces that have been *averaged* together. When such composites, or averaged faces are presented to observers, they are typically perceived as more attractive than are the individual faces that went into the composite. Figure 1(a) shows an averaged female face created by combining a number of individual faces.

Why is this the case? Some argue that facial composites represent the population average—which is likely to be processed by observers with great ease. Others argue that facial composites are deemed attractive because they enhance visual cues that are theoretically related to health and fertility: symmetry and some aspects of neoteny such as smoothness of skin. Regardless of the underpinnings, “average” faces are indeed attractive.

Cues to Gender Typicality/Sexual Dimorphism

Gender typicality emerges as an important factor for both faces and bodies, and extremity of

Figure 1 Two "Averaged" Faces Created by Computer Averaging of a Large Number of Individual Faces

Source: Courtesy of Ben Jones and Lisa DeBruine, Face Research Lab, University of Aberdeen.

gendered cues can, under some circumstances, be an important cue to attractiveness. In face perception research, results differ for male and female faces. Feminine female faces (or gender typical faces) are consistently perceived to be attractive by men and women alike. Perceptions of attractiveness for male face, in contrast, show a different pattern. Specifically, the extent to which masculine facial morphology is perceived to be attractive varies across women's menstrual cycle. At the peak of the fertile window of their cycle, women tend to find more masculine male faces to be attractive; at other periods, women find more feminine male faces to be attractive. Because facial masculinity has been linked to health, it is argued that women's cyclical change in preferences reveals women's desire to find mates with good genes.

In body perception research, early reports indicated that men preferred women with hourglass figures and argued that such preferences were adaptive because body shape relates to health and fecundity. Subsequent research challenged this interpretation, however. Men's preferences for specific body shapes were found to vary dramatically across cultures, and the health outcomes indicated tend to occur late in life (i.e., after childbearing years). Moreover, other research demonstrated that the perception of gender typicality, rather than perceptions of health or fertility, was a stronger predictor of perceived attractiveness. Bodies that combined female typical shape with feminine motion or male typical shape with masculine motion, for example, were perceived to be attractive. Other combinations were not. Thus, bodies are perceived to be attractive when they are gender typical. In some cases, the most attractive cues are quite extreme.

Elected Cues

Much of the prior research that specified cues related to perceived attractiveness has focused on those that are difficult to change (i.e., biologically determined). Without surgical intervention, for example, it is difficult to alter one's facial symmetry or the body's structure. Other cues, however, are more easily modified. Manner of dress, application of cosmetics, and physical adornment (e.g., tattoos and piercings) are *elected* or chosen by individuals. Although the specific manifestation of such cues may vary, the overall tendency reveals universal themes. The nature and location of adornment tends to accentuate the factors discussed. Cosmetic application, for example, is more common in women and tends to accentuate neotenous features by changing the apparent size of the eyes. Clothing styles and body adornment such as tattoos tend to accentuate sexually dimorphic regions of the body (shoulder breadth for men, and breasts and a small waist for women). Research has even demonstrated that at peak fertility, women tend to expose more skin and dress in a slightly more attractive manner.

Auditory Cues to Attractiveness

Like visual cues in the face and the body, voice cues also carry important information that affects perceptions of attractiveness. Many such cues convey sexual dimorphism, health, and fertility. Testosterone exposure during puberty alters young men's vocal folds thus changing voice's fundamental frequency (the acoustic correlate of perceptual pitch). Voice pitch is one sexually dimorphic cue that affects the perceived attractiveness of both men and women. More masculine voices (that have a lower pitch) are generally perceived to be more attractive for men. This tendency is more pronounced when female raters are at or near peak fertility. Similarly, more feminine voices (that have a higher pitch) are generally perceived to be more attractive for women. Moreover, women's voices are evaluated to be most attractive when they are at the most fertile point in their cycles. Finally, some evidence indicates that the perception of attractiveness in voices covaries with bilateral symmetry of the body, an indicator of genetic fitness.

Olfactory Cues to Attractiveness

Scent proves to be a multimillion-dollar industry annually, so it is not surprising that scent is a potent cue to attractiveness. Research in the attractiveness of scent, however, has focused not on people's preferences for perfume, but on their evaluations of body odors. When it comes to scent, beauty really is in the eye (or the nose) of the beholder.

Research suggests that the body odor that one finds attractive is highly systematic, though at times idiosyncratic. Several distinct effects support this claim. First, both men and women appear to prefer the body odors of members of the opposite sex who have *dissimilar* immunities (produced by major histocompatibility complex genes). According to this research, such preferences make sense insofar as they increase the likelihood that offspring will benefit from the diversity of genes from mom and dad, receiving "the best of both worlds," or in this case, the best from both parents. Second, scent preferences fluctuate with fertility (for both men and women). Men, for example, appear to prefer the body odors of women who are at or near peak fertility. This tendency, it is argued, promotes sexual attraction and interaction during the narrow window in which a woman can conceive. Women reliably prefer the aroma of symmetrical men (linked to genetic fitness), but only near ovulation when the risk of conception is high. Thus, olfactory cues to attractiveness appear to solve adaptive problems, and the particular scent that is deemed attractive is likely to vary across time and individual.

Perceptions of attractiveness are multifaceted, supported by numerous visual, auditory, and olfactory cues. Many such findings have been described as evolved preferences that serve adaptive ends. Regardless of their origins, cues to attractiveness show remarkable stability across age and culture.

Kerri Johnson

See also Motion Perception: Social; Social Perception

Further Readings

Etcoff, N. (1999). *Survival of the prettiest: The science of beauty.* New York: Doubleday.

Haselton, M. G., Mortezaie, M., Pillsworth, E. G., Bleske-Recheck, A. E., & Frederick, D. A. (2007). Ovulation and human female ornamentation: Near ovulation, women dress to impress. *Hormones and Behavior, 51,* 40–45.

Johnson, K. L., & Tassinary, L. G. (2007). Compatibility of basic social perceptions determines perceived attractiveness. *Proceedings of the National Academy of Sciences USA, 104,* 5246—5251.

Rhodes, G., & Zebrowitz, L. A. (Eds.). (2002). *Facial attractiveness: Evolutionary, cognitive, and social perspectives.* Westport, CT: Ablex.

AUDIOLOGY

The number of hearing-impaired people worldwide is estimated to be around 300 million and rising because of the growing global population and longer life expectancy. Hearing impairment is now the third leading chronic disability after arthritis and hypertension in the Western world. Therefore, most individuals, or members of their family, will attend an audiology service to have their hearing assessed at some point in their lives. *Audiology* is the science of hearing, but the term is used in a healthcare setting to generally mean the study and assessment of hearing and balance problems and the treatment and prevention of disorders of these functions. This entry describes various audiological assessment procedures.

Hearing Sensitivity

The human ear has an extremely wide perceptual dynamic range. The lower limit of hearing where sound is just detectable is referred to as the *threshold of hearing* (also known as absolute threshold or absolute sensitivity). The upper limit of hearing where sound begins to become uncomfortably loud is referred to as the *threshold of discomfort* (or uncomfortable loudness level). In quantitative terms, the difference in level between these two extremes is of the order of 10 million times. The human ear can hear single frequencies of vibration from around 20 to 20,000 hertz (Hz), although the upper limit, in particular, reduces during the natural ageing process.

The general relationship between the dynamic range of hearing and frequency has been well understood for many years. Studies measuring the minimum audible level of hearing have been made with stimuli presented to each ear separately or both ears together (usually via an earphone and loudspeaker, respectively). The results from the two methods are similar, but not identical, and show human hearing to be generally most sensitive between 500 and 10,000 Hz. The typical values obtained in a group of young healthy individuals, at individual frequencies, are used as the baseline reference level to which listeners with a suspected hearing impairment can be compared.

The Pure Tone Audiogram

The most widely used assessment procedure in clinical audiology is known as pure tone audiometry. The listener's hearing threshold level, in decibels (dB), is plotted on a chart, known as a pure tone audiogram, with hearing threshold level plotted on the ordinate as a function of signal frequency on the abscissa. The conventional clinical audiogram plots hearing level with low values (normal hearing) at the top of the chart and raised levels closer to the abscissa. Therefore, raised hearing levels are plotted lower on the pure tone audiogram chart (see Figure 1, next page).

The reference baseline is called audiometric zero and represents the zero decibels hearing level line on the audiogram chart. If, for example, a listener's hearing threshold level for a particular signal is 60 dB, this means that the listener's hearing threshold is 60 dB higher than the median value obtained from a group of normal healthy young persons. Not every healthy young adult has a hearing threshold level that falls on the zero decibels line, however, and the range of normality is generally taken to be ±20 dB of the zero line.

For clinical purposes, the hearing threshold is usually measured for single frequency tones at discrete frequencies from 500 Hz to 8,000 Hz, in octave or half-octave intervals, and reported in step sizes of 5 dB. The signals are selected and presented to the listener using a classical measurement method known as the methods of limits. A series of signals are presented in an ascending or descending run (from loud to quiet, or vice versa), and the task for the listener is to respond every time he or she detects a signal. As with any psychophysical measurement, there will be a level above which the pure tone is always heard and a lower level where the tone is never heard. The threshold of hearing is taken as the lowest level at which the signal is detected at least 50% of the time. A whole host of extrinsic and intrinsic factors can influence the measurements (e.g., ambient noise level and duration of the test signal, respectively), so there are clearly defined procedures for clinical testing.

Although the measurement of hearing thresholds from each ear separately appears a relatively straightforward procedure, sound may cross from one side of the head to the other. For example, if a listener has poor hearing in the ear under test, the signal may be sufficiently intense that it may cross the skull and be detected by the opposite ear that has better hearing. In these circumstances, a masking noise is presented to the non-test ear to prevent *cross hearing* of the

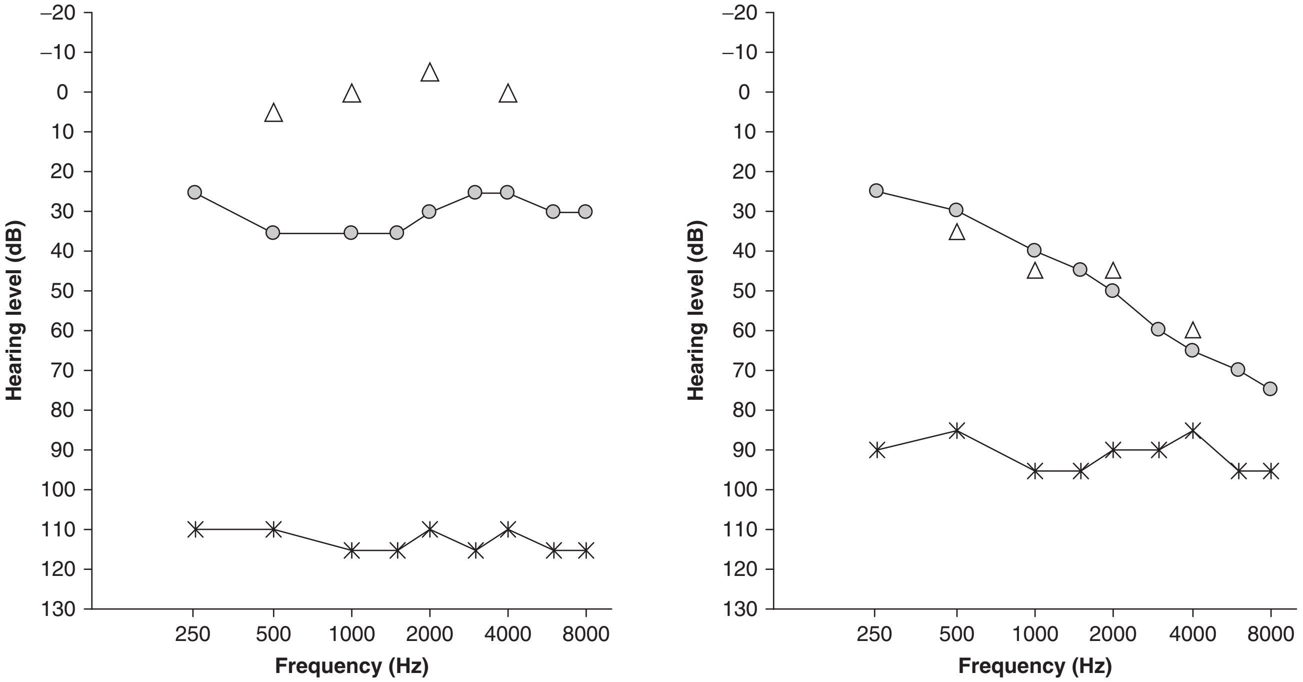


Figure 1 Pure Tone Audiometry Data for a Listener With a Mild Hearing Impairment (left panel) and a Listener With a Moderate High-Frequency Sloping Hearing Impairment (right panel)

Notes: Only data from one ear of each listener is shown: right air conduction, circles; right bone conduction, triangles; loudness discomfort level, asterisks. The presence of an air-bone gap in the left panel means the listener has a conductive hearing impairment. The lack of an air-bone gap on the right panel means that this listener has a sensorineural hearing impairment. The ear with the conductive impairment has a normal dynamic range, that is, air conduction thresholds and loudness discomfort levels are both elevated (lower on the chart) by approximately 30 decibels compared with that of a normal listener. The loudness discomfort levels in the right panel are present at normal levels, despite the raised hearing thresholds: this listener has a greatly reduced dynamic range of hearing. The findings in the left panel are typical of an individual who has fluid in the middle ear space. The findings in the right panel are typical of a listener with a natural age-related hearing impairment.

test signal. Standard procedures have been developed for when and how to use masking. If masking is insufficient, the test signal may continue to be detected by the non-test ear: if too much masking is used, this may result in *cross masking* and artificially raise the hearing threshold in the test ear.

The degree of hearing impairment can, and usually does, vary with frequency although it is usually summarized as mild, moderate, severe, or profound, based on an average of the hearing threshold levels. The ability to hear speech is related to the degree of impairment. Slight hearing impairment (26–40 dB hearing level) can cause some difficulty following speech, especially in noisy situations. Moderate hearing impairment (41–60 dB hearing level) can cause difficulty following speech without a hearing aid. Conversational speech will not be audible in cases of severe (61–80 dB hearing level) or profound (81 dB hearing level or greater) hearing impairment, and some of these listeners may need a special type of hearing aid known as a cochlear implant if they are to perceive speech. The proportion of speech that is audible and useable for a listener, with or without a hearing aid, can be quantified using a procedure known as the Speech Intelligibility Index. At relatively high presentation levels, the test signal can sometimes be perceived as a vibration, especially for low frequency stimuli. Therefore, sometimes it can be difficult to determine if a threshold measurement is an auditory or a vibro-tactile perception.

Conductive and Sensorineural Hearing Impairment

Hearing impairment is generally categorized into two groups: conductive and sensorineural. Conductive hearing impairment occurs when there is a problem in the outer or middle ear that prevents sounds being conducted to the cochlea in the inner ear. Sensorineural hearing impairment involves a problem with either the *sensory* transducer cells in the cochlea or, less commonly, the *neural* pathway to the brain. Conductive hearing-impairment can often be corrected via surgery and is relatively common in childhood, but sensorineural hearing impairment is usually permanent. Therefore, it is important to distinguish between the two categories. One method of doing this is to compare air conduction and bone conduction hearing threshold levels. This involves measuring hearing sensitivity using two different types of earphone. When a pure tone is presented via an earphone (or a loudspeaker), the signal travels through the air in the outer ear to the middle ear and then to the cochlea in the inner ear. This is known as *air conduction testing*. Alternatively, the outer and middle ear can be largely bypassed by stimulating the cochlea via mechanical vibration of the skull. This is known as bone conduction testing. Instead of using an earphone, an electromechanical earphone is placed on the skull.

Normal hearing individuals will have a hearing threshold level close to zero decibels for both air and bone conduction. Disorders of any part of the auditory pathway will affect the air conduction threshold, but disorders of the conducting mechanism will have much less effect on bone conduction measurements because these generally bypass the outer and middle ear. When a hearing impairment is present but the air-bone gap (air conduction minus bone conduction) is close to zero, it is assumed that there is no impairment of the outer or middle ear and the listener has a sensorineural hearing impairment. The presence of an air-bone gap signifies a conductive hearing impairment.

The dynamic range between the threshold of hearing and loudness discomfort level is around 80 to 100 dB in normal hearing listeners. Listeners having a sensory hearing-impairment will have raised hearing thresholds, but their loudness discomfort levels are essentially similar to those of normal hearing listeners. Listeners with a sensory hearing impairment have a reduced dynamic range and an abnormal rate of loudness growth: This is known as loudness recruitment, that is, an abnormally disproportionate increase in loudness for a small increase in intensity. This has implications for the design of hearing instruments because soft sounds will require greater amplification than will loud sounds—that is, nonlinear amplification is required. Although a nonlinear hearing instrument can compensate by providing more amplification for soft sounds, it cannot compensate for the loss of supra-threshold abilities such as impaired frequency resolution. This means that hearing in background noise remains a problem for many listeners.

Pediatric Assessment Procedures

Between a developmental age of 6 months and 3 years, the measurement technique of choice is visual reinforcement audiometry. This involves pairing a head turn response to a sound with an interesting visual reward such as a flashing light or an animated toy animal. Once this classical conditioning has been established, operant conditioning then takes place where a visual reward is presented after an appropriate sound-elicited head turn. This technique is used to determine the minimum response level that will elicit a head turn. Although it is usual to attempt ear-specific measurements in children, in some cases, earphones will not be tolerated and the signal is presented from a loudspeaker: this is known as sound field audiometry. Hearing sensitivity in infants is slightly raised compared with that of young adults although there is debate about whether this is purely a physiological maturation or if it is related to nonsensory factors such as attention and motivation.

Before a developmental age of 6 months, behavioral testing is of limited use in determining hearing threshold levels. However, a small amount of sound is generated in the healthy cochlea, and this otoacoustic emission can be measured with a small sensitive microphone in the ear canal. This normal response from a healthy ear forms the basis of a clinical procedure that can be used, for example, to screen the hearing of a newborn baby. If no otoacoustic emission can be recorded, event-related potentials can then be used to estimate hearing sensitivity. This involves the measurement of electrical potentials via recording leads attached to the scalp. The method of choice in infants is the auditory brainstem response because this can be obtained during sleep. A typical procedure is to commence at a high level and reduce this until the evoked response can no longer be detected. The presence of a response is based on the tester's subjective interpretation of the waveform. An alternative technique that is gaining popularity is the auditory steady state response, and the presence of this response is based on statistical data. Event-related potentials can also be used to estimate hearing sensitivity in adults who are unable or unwilling to provide reliable information via pure tone audiometry. Historically, event-related potentials have had a valuable diagnostic role, especially when identifying tumors associated with the hearing nerve, but these have largely been replaced by advanced diagnostic imaging techniques.

A commonly used procedure that provides information about the condition of the eardrum and the middle ear is known as tympanometry. Although this can also be used in adults, it is particularly useful in children who are prone to fluid gathering in the normally air-filled middle ear space. The procedure works on the principle that some sound entering the ear canal is reflected back from the eardrum and this can be measured with a sensitive microphone. Stiffening the eardrum by changing the pressure of the air trapped in the ear canal should result in more sound being reflected by the eardrum. However, if there is fluid in the middle ear, then the stiffness of the ear drum will be unaffected by changing the air pressure in the ear canal.

Vestibular Assessment and Management

Sudden changes in the function of the vestibular organ in the inner ear can result in rotatory vertigo, which gives the illusion that the environment is spinning around. Useful information about vestibular function can be obtained by observing eye movements during certain visual and vestibular stimulation. The audiologist is particularly interested in the presence of a slow-quick oscillatory movement of the eyes known as nystagmus. This eye movement will be present spontaneously after a change in vestibular function and may continue for days or weeks until the brain has time to compensate. Nystagmus may also be provoked by changes in body position such as getting out of bed in the morning. The sensitivity of the right and left vestibular organs can be compared by irrigating the external ear canal with hot or cold water to induce a response: This is known as the caloric test. Increasingly, audiologists are using a force platform to measure body sway because this can provide information about the use of the visual, vestibular, and proprioceptive systems for balance function and postural control. Rehabilitative procedures generally involve head and eye exercises that aid the central compensation mechanism.

Kevin J. Munro

See also Ageing and Hearing; Audition: Disorders; Auditory Thresholds; Cochlear Implants: Technology; Evoked Potential: Audition; Hearing Aids; Tinnitus; Vestibular System

Further Readings

British Society of Audiology (2008, March). Recommended procedure: Pure tone air and bone conduction threshold audiometry with and without masking and determination of uncomfortable loudness level. Retrieved from http://www.thebsa.org.uk/docs/RecPro/PTA.pdf

Gelfand, S. A. (2004). *Hearing: An introduction to psychological and physiological acoustics* (4th ed.). New York: Marcel Dekker.

Haughton, P. (2002). *Acoustics for audiologists.* Oxford, UK: Academic Press.

Katz, J. (2002). *Handbook of clinical audiology* (5th ed.). London: Lippincott, Williams and Wilkins.

Moore, B. C. J. (2007). *Cochlear hearing loss: Physiology, psychology and technical issues* (2nd ed.). London: Wiley Blackwell.

Newborn Hearing Screening Progamme (2008, March). Audiology protocols for newborn hearing screening: Visual reinforcement Audiometry testing of infants. Retrieved from http://hearing.screening.nhs.uk/cms.php?folder=21

Plack, C. J. (2005). *The sense of hearing.* Mahwah, NJ: Lawrence Erlbaum.

World Health Organization (2008, March). Factsheet: Deafness and hearing impairment. Retrieved from http://www.who.int/mediacentre/factsheets/fs300/en/index.html

AUDITION

It seems likely that *audition,* commonly known as hearing, initially evolved to alert organisms to significant events in the environment. Unlike visual stimuli, sounds can be detected whatever their direction of incidence. Hearing can indicate the presence of a predator or prey, and it can indicate the appropriate location to direct visual attention. Hearing can also convey information about the properties of sound-emitting objects, such as their size and whether they are approaching or receding, which may determine whether they are threatening or not. In many species, and especially in humans, hearing has evolved further as a means of communication, speech communication being the prime example. Indeed, the ability to convey information using sounds as symbols for objects and actions is one of the characteristics that distinguishes humans from other animals, as this entry describes.

How Information Is Carried in the Auditory System

Sounds are transmitted through the outer ear and middle ear into the inner ear or cochlea, where they are subjected to a frequency analysis on the basilar membrane; each place is tuned to respond best to a limited range of frequencies. Low frequencies produce their biggest response toward the apex of the cochlea, and high frequencies produce their biggest response toward the base. Each neuron in the auditory nerve derives its response from the vibration at a specific place on the basilar membrane.

Information about the characteristics of sounds is carried in the auditory nerve in three basic ways:

1. By the rate of firing of individual neurons, which will be referred to as the "amount" of neural activity. The more vibration there is at a given place, the greater is the amount of activity in neurons connected to that place. It is commonly believed that the subjective loudness of a sound is related to the amount of neural activity evoked by that sound, although this idea has been disputed.

2. By the distribution of activity across neurons. Each neuron is tuned so that it responds most strongly to a specific frequency, called the characteristic frequency (CF); the tuning reflects the tuning of the place on the basilar membrane that the neuron is driven by. The distribution of amount of neural activity as a function of CF is called the *excitation pattern.* The concept of the excitation pattern, developed by pioneers such as Harvey Fletcher and Eberhard Zwicker, plays an important role in many theories of auditory perception—for example, theories of the perception of pitch, loudness, and timbre. The excitation pattern can be considered as conveying "place" information because the CF at the peak of the excitation pattern is related to the place on the basilar membrane that is excited most.

3. By the detailed time-pattern of the neural impulses, and especially the time intervals between successive nerve impulses. This form of information is known as "temporal" information. Neural impulses tend to be evoked at times corresponding to a specific phase of the waveform on the basilar membrane, an effect called *phase locking.* As a result, for a periodic sound, the time intervals between successive nerve impulses are approximately integer multiples of the period of the sound. Phase locking breaks down at high frequencies (above about 3–6 kilohertz [kHz] in most mammals), but the upper limit in humans is not definitely known. Studies of pitch perception suggest that the limit in humans is about 5 kHz.

In addition, information about sounds is conveyed by the differences between the two ears in all of those methods previously mentioned. In particular, differences in intensity at the two ears (which are primarily conveyed by differences in neural firing rate) and differences in the time of arrival of sounds at the two ears (conveyed mainly by subtle differences in the exact timing of nerve spikes at the two ears) play a strong role in determining the perceived location of sounds.

Knowledge about the way that the properties of sounds are represented in the auditory nerve has played an important role in the design of cochlear implants, which are used to restore hearing to people with severe to total hearing loss; these implants restore hearing by electrical stimulation of the auditory nerve via an electrode array implanted into the cochlea. The "coding schemes" used in cochlear implants attempt to mimic, albeit in a crude way, the neural patterns that would be evoked by sound in the auditory nerve of a person with normal hearing.

Place Versus Temporal Information

A classic debate in hearing theory is concerned with the relative importance of place and temporal information. Originally, this debate revolved mainly around the relative role of place and temporal information in the perception of the pitch of pure and complex tones. The debate about pitch continues, but most researchers now believe that pitch perception involves the use of both place and temporal information; indeed, there might have to be a correspondence between the two types of information before a clear pitch can be perceived. More recently, researchers have proposed that both place and temporal information play a role in several other aspects of auditory perception, including masking (the process whereby one sound is rendered inaudible by the presence of another sound) and speech perception.

Temporal information may be especially important in situations where a background sound is fluctuating in level from moment to moment. The normal auditory system can make use of information present during brief dips in a background sound to enhance perception of a target sound. This is often called "listening in the dips." For example, a target talker can be understood much better in a fluctuating background sound than in a steady background sound with the same average level. This ability seems to depend at least partly on the use of temporal information; changes in the temporal information during the dips can be used to determine that a signal of interest is present, and what its characteristics are. Interestingly, recent data suggest that even relatively mild hearing loss can adversely affect the ability to use temporal information, and this may help explain why hearing-impaired people usually get little benefit from dip listening.

Peripheral Versus Central Processes

Peripheral processes in hearing are those occurring in the outer and middle ear, in the cochlea, and in the auditory nerve. Central processes are those occurring at stages of the auditory system following the auditory nerve. Many theories and models of auditory perception are based on relatively peripheral processes. For example, models of loudness perception, such as developed by Harvey Fletcher, Eberhard Zwicker, Brian Moore, and their coworkers, consider the effect of the outer and middle ear on the sound transmitted to the cochlea, and also consider the processing of sound within the cochlea. In particular, the influence of the "active mechanism" in the cochlea, which introduces strong nonlinear effects, has been studied intensively. However, processing occurring in the central auditory system has been largely ignored. Early theories of pitch perception were also largely based on the information that was present in the auditory nerve. Only in the case of sound localization, which requires comparison of the neural signals from the two ears, were central processes considered. The trend during the last 20 years has been to consider the role of central processes in much more detail, both at the physiological and the psychological level. The trend has partly come from an increasing realization of the limits of explanations based on peripheral processes. It makes sense to consider the role of peripheral processes and the extent to which they can explain the perceptual phenomenon in question, but once such explanations have been exhausted, more central processes have to be considered.

Energetic and Informational Masking

An example of the increasing consideration of central processes comes from studies of auditory masking, which is the process by which one sound (the masker) makes it difficult or impossible to hear another sound (the signal or target). Many of the cases of masking studied in the laboratory, for example, the detection of sinusoidal tones in white noise, can be explained largely in terms of processes occurring in the cochlea and auditory nerve. Some models of masking can predict the detection thresholds of typical human listeners for such cases with high accuracy. Indeed, these models are incorporated in digital systems for storing and transmitting sounds, such as MP3 and the systems used in transmitting cell (mobile) telephone signals; such systems are referred to as perceptual coders because they exploit the (predictable) properties of masking in the human auditory system to reduce the number of "bits" needed to transmit or store the audio signal. Masking of this type is sometimes called (inappropriately) energetic masking, and it occurs when the response of the auditory nerve to the masker-plus-signal is similar to the response to the masker alone. In other words, the signal is masked because the information conveyed by the auditory nerve is inadequate for its detection.

When a masking sound is highly similar in some way to the signal, or when the properties of the masker vary in an unpredictable way from one stimulus to the next, there may be much more masking than would be expected from energetic masking alone. This "extra" masking is called informational masking. It is assumed that informational masking occurs because the signal is confused with the masker, or because attention is directed to an inappropriate aspect of the sound.

When the task of a listener is to identify the speech of one talker (the target) in the presence of another talker (the background) with similar characteristics, informational masking may play a strong role. Under these conditions, the amount of informational masking is greatly influenced by whether or not the target and background are perceived to come from the same location in space. Richard Freyman and his coworkers illustrated this in an experiment. Listeners were asked to identify nonsense syllables spoken by a female talker in the presence of a background of either noise (with the same average spectrum as the speech) or a second female talker. In the case where the background was a talker, it was difficult to determine which parts of the sound came from the target talker and which came from the background because the two voices were similar in quality. In one pair of conditions, the target and background were both presented from two loudspeakers, one located directly in front of the listener and one located 60° to the right. The target speech from the loudspeaker at 0° was presented slightly earlier in time than the target speech from the loudspeaker at 60°, which made the target speech appear to come from in front. This happens because of an effect called the precedence effect. In one condition (called the coincident condition), the background sound was presented in the same way, so that it too was heard as coming from in front. In a second condition (called the separated condition), the background sound was presented slightly earlier in time from the loudspeaker at 60° than from the loudspeaker at 0°, which made the background appear to come from the right. For the noise background, the percentage of key words identified was only 5 to 10% better for the separated than for the coincident condition. However, for the female-talker background, the corresponding difference was 15 to 30%. The greater effect for the speech background was probably caused by a release from informational masking caused by the perceived spatial separation of the target speech and the background. When the target was perceived as spatially separated from the background, it was easier to determine which parts of the sound came from the target talker.

Informational masking, and the conditions under which it occurs, has been an area of increasing research activity during the last 20 years. One finding of this research is large individual differences; some people appear to be much more susceptible to informational masking than are others. There can also be considerable learning effects, with informational masking decreasing as training proceeds over many weeks of daily training. The existence of such large learning effects confirms that the masking is not determined by peripheral processes, but depends on limitations in central mechanisms, which can partly be overcome by training.

Veridical Perception and Auditory Illusions

Following pioneers such as Hermann von Helmholtz, many researchers in the field of perception take the view that our sensory systems are designed to help us determine the true nature of the outside world. When such veridical perception fails, the percept is often described as an illusion. This view is more difficult to assess experimentally for hearing than for vision. For vision, it is usually relatively obvious what the "correct" perception ought to be; for example, a straight line ought to be perceived as straight. For hearing, it is often much less obvious how a sound ought to be perceived, and it is often not clear whether an auditory percept should be described as an illusion.

In the early days of research on the perception of the pitch of complex sounds, the phenomenon of the missing fundamental—the perception of a pitch corresponding to the fundamental frequency, even when the sound contains no component at the fundamental frequency—was sometimes described as an illusion. Nowadays, the phenomenon is simply assumed to reflect the way that pitch perception normally works. The pitch of a complex tone is usually determined by components (higher harmonics) other than the fundamental frequency. It became inappropriate to refer to the phenomenon as an illusion once better theories of pitch had been developed.

As another example, consider the so-called continuity illusion. When a sound A is alternated with a sound B, and B is more intense than A, then A may be heard as continuous, even though it is interrupted. Quite complex aspects of the percept may be heard as continuing through the interruption. For example, if B is noise and A is a tone that is gliding upward in frequency, the glide is heard as smooth and continuous even though certain parts of the glide are missing. If A is a tone that is amplitude or frequency modulated, so that it sounds like it is fluctuating in loudness or pitch, the fluctuation is heard to continue during the time that the tone is replaced by noise. Notice that, for the perceived continuity to occur, the gaps in the tone must be filled with noise and the noise must be a potential masker of the tone (if they were presented simultaneously). In the absence of noise, gaps in the tone are clearly heard.

Although the continuity effect is described as an illusion, the auditory system's interpretation of the tone A as being continuous is reasonable. If the tone were continuous, the pattern of neural activity evoked by the tone plus the noise bursts would not be detectably different from the pattern evoked by the alternating tone and noise. Generally, the auditory system seems to create a percept that corresponds to the most likely or most plausible interpretation of the sensory information. It is thus not clear whether it is appropriate to describe the continuity effect as an illusion.

Auditory Scene Analysis and the Concept of Auditory Events and Objects

The sound reaching our ears rarely comes from a single source. Rather, several sources are often active simultaneously. The auditory system is faced with the task of analyzing the complex mixture and deriving a perceptual representation of each individual sound source. This task has been described as being analogous to monitoring the pattern of waves at two closely spaced points on the surface of the sea (corresponding to the acoustic waves at the two eardrums) and using this information to work out the nature and position of all of the ships, boats, and swimmers in the vicinity. The task is immensely difficult and complex, and most computational methods that have been developed to perform the task—often described using the term *computational auditory scene analysis*—perform much more poorly than do human listeners.

Albert Bregman has emphasized the distinction between two concepts: source and stream. A sound source is some physical entity that gives rise to acoustic waves, for example, a person talking. A stream is the percept of a group of successive or simultaneous sound elements as a coherent whole, appearing to come from a single source. For example, it is the percept of hearing a person talking. Some researchers have argued that the auditory system has evolved primarily to allow it to accomplish the task of deriving streams when multiple sources are active. Millions of years of evolution have led to the superiority of human listeners over machines in performing this task.

In recent years, the phrase *auditory object* has often been used as an alternative to *stream*. Also, sometimes the term *auditory event* is used. However, there is no universal agreement about what is meant by an auditory object or event, apart from a general consensus that both are perceptual entities. The difficulty can be illustrated by considering the situation of listening to a violin being played, or a person talking. In each case, the source produces a sequence of sounds (complex tones or words), which are usually not separated by silences. The perceptual experience is that of hearing a sequence of sounds. The auditory system somehow derives discrete perceptual representations of the individual notes or words despite the continuous nature of the sound wave. Each of the perceived sounds (each note or word) may appear different from the preceding and following sounds, but, at the same time, the sounds from a given source appear coherent; they all have the characteristic quality of that specific violin (as played by that specific player) or that specific talker.

The questions thus arise: Does each of the perceived sounds correspond to an auditory object, or is the object more appropriately defined in terms of the perception of the sequence of sounds? If the former, can a single note played on a violin, or a single word uttered by a talker, be perceived as an auditory object? Does the "object" correspond to the individual note or word, or to the source of the note or word? The following classification scheme may be useful in answering these questions.

Any sound source can evoke perceptual attributes, such as loudness, pitch, timbre, roughness (which is linked to rapid fluctuations in amplitude), and apparent location. When these attributes are perceptually linked, as if they all belong together, then the joint percept may be described as an auditory event. When that auditory event is associated with a particular sound source, that is, the event is perceived as being produced by the source, the perceived sound source may be described as an auditory object. According to this classification scheme, a sound such as a single note or word can (and usually does) lead to both an auditory event and an auditory object. The auditory event corresponds to the perceived note (with its specific pitch, timbre and loudness) or the perceived word (with its specific pitch, timbre, loudness, and identity), whereas the auditory object corresponds to the perceived sound source, the violin or the talker.

Perceptual Constancy

Perceptual constancy refers to the fact that the perceived properties of objects remain relatively constant despite changes in the conditions of viewing or listening. For example, the shape of a coin does not appear to change when it is viewed from an oblique angle. Perceptual constancy has been studied intensively for the visual system, but has received relatively little attention for the auditory system. Consider the case of loudness perception. The physical intensity of the sound reaching a listener's ears depends partly on the sound intensity emitted by the source and partly on the distance of the listener from the source. It remains somewhat unclear whether loudness is related more to the intensity at the ears or to the intensity of the sound emitted by the source. However, at least when the sound source is visible, a form of constancy sometimes holds; the intensity at the source plays a stronger role than does the intensity at the ears. In other words, the auditory system appears to base loudness mainly on the properties of the source, the distal stimuli, rather than on the properties of the sound reaching the ears (the proximal stimulus). For example, when the sound intensity at the ears is held constant, judgments of loudness are influenced by the perceived distance of the source, as determined visually.

Perceptual constancy may also play a role in the perception of the timbre of sounds, which refers to their characteristic quality. If two tones have the same loudness and pitch but are perceived as different, then the difference is described as a difference in timbre. Timbre depends partly on the shape of the spectrum of sounds. In many everyday situations, the sound reaching the ears is composed partly of sound coming directly from the source and partly of sound that has been reflected from nearby surfaces and reaches the ears after a delay. The interference of the direct and reflected sound changes the spectrum of the sound at the ears, relative to the sound emitted by the source. The influence of the reflected sound has been studied extensively for its effects on sound localization, which are surprisingly small. However, the effect on timbre perception has received little attention. One might expect the reflected sound to have a strong influence on perceived timbre, but everyday experience suggests that this is not the case. Rather, perceptual constancy seems to operate. For example, the sound quality of a familiar person's voice does not seem to differ markedly when listening to that person outdoors and or a reverberant room. Studies conducted by Tony Watkins and colleagues suggest that listeners are able to compensate for the effects of room acoustics (reflections from the walls, floor, and ceiling of a room), at least for identifying speech sounds. However, further work in this area is needed.

Brian C. J. Moore

See also Audition: Disorders; Audition: Loudness; Audition: Pitch Perception; Auditory Frequency Analysis, Neural; Auditory Illusions; Auditory Localization: Physiology; Auditory Localization: Psychophysics; Auditory Masking; Auditory Processing: Central; Auditory Processing: Peripheral; Auditory Receptors and Transduction; Auditory Scene Analysis; Auditory System: Evolution of; Constancy; Speech Perception; Visual Masking

Further Readings

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge: Bradford Books, MIT Press.

Darwin, C. J., & Carlyon, R. P. (1995). Auditory grouping. In B. C. J. Moore (Ed.), *Hearing* (pp. 387–424). San Diego, CA: Academic Press.

Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *Journal of the Acoustical Society of America, 106,* 3578–3588.

Kidd, G. R., Watson, C. S., & Gygi, B. (2007). Individual differences in auditory abilities. *Journal of the Acoustical Society of America, 122,* 418–435.

Litovsky, R. Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. (1999). The precedence effect. *Journal of the Acoustical Society of America, 106,* 1633–1654.

Mershon, D. H., Desaulniers, D. H., Kiefer, S. A., Amerson, T. L., & Mills, J. T. (1981). Perceived loudness and visually-determined auditory distance. *Perception, 10,* 531–543.

Moore, B. C. J. (2003). Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants. *Otology & Neurotology, 24,* 243–254.

Moore, B. C. J. (2003). *An introduction to the psychology of hearing* (5th ed.). San Diego, CA: Academic Press.

Moore, B. C. J. (2007). *Cochlear hearing loss: Physiological, psychological and technical issues* (2nd ed.). Chichester, UK: Wiley.

Moore, B. C. J., Glasberg, B. R., & Baer, T. (1997). A model for the prediction of thresholds, loudness and partial loudness. *Journal of the Audio Engineering Society, 45,* 224–240.

Watkins, A. J. (1991). Central, auditory mechanisms of perceptual compensation for spectral-envelope distortion. *Journal of the Acoustical Society of America, 90,* 2942–2955.

Zahorik, P., & Wightman, F. L. (2001). Loudness constancy with varying sound source distance. *Nature Neuroscience, 4,* 78–83.

AUDITION: COGNITIVE INFLUENCES

Audition is the ability to sense and perceive sound. Audition relies on bottom-up sensory mechanisms, which govern how the sound is changed from physical energy in the environment to electrical impulses used by the brain, and top-down cognitive perceptual mechanisms, which govern how we identify and interpret sounds. The ability to transduce sound is only one small part of audition because the cognitive abilities we bring to bear when listening are also important for auditory perception. For example, imagine that you are walking down a dark street at night and you hear the sound of a dog barking. A wealth of auditory information is available in the signal that will tell you about the dog and determine how you will react to it. Bottom-up processes will allow you to identify where the dog is located (in front or behind you; on your left or right), tell you about its proximity (near or far away), and whether the dog is outside in a yard (if the sound is clear and unobstructed) or inside a building (if the sound is muffled and filtered by the walls). However, top-down mechanisms will influence your perception of the auditory information and determine how you will react to the sound. Such mechanisms allow you to accurately identify the sound as being produced by a dog rather than some other source (a cat, or a person imitating a dog) and allow you to determine its size (small or large) or possibly its breed (such as a Bassett hound, Chihuahua, or Rottweiler). You may also be able to determine the mood of the dog (does it sound friendly or unfriendly? does it sound excited, sad, fearful, or aggressive?). If you have had bad experiences with dogs, you may interpret the sounds as being aggressive, eliciting a fear response. If you have had good experiences with dogs, the sound may elicit a more neutral or positive response. Even in the seemingly simple case of hearing a barking dog, a variety of cognitive processes are brought to bear when you interpret the auditory information and will ultimately determine your response. This entry describes adaptation, attention, and context of cognitive influences on audition, as well as perceptual learning and expertise.

Adaptation, Attention, and Context

From the previous example, we can see that top-down information can alter our perception of sound. The knowledge that a listener has about sound and how the listener uses that information is an important part of audition. The amount of attention that is required during listening and the context of the environment play central roles in altering our perception of sound. For example, you quickly adapt to the din of the conversations around you when in a noisy restaurant: Instead of hearing individual voices, you hear a rumbling babble in the background. Yet, occasionally a sound will emerge and become perceptually salient (such as when someone coughs loudly, or we hear our name being called). Adaptation to an environmental context will allow us to move our focus of attention from things that do not concern us (such as the conversation at the next table) to things that are more important (such as the other people at our table or hearing our own name). The ability to separate sound sources and focus attention on a particular source (in this case a specific talker) is central to our success in listening in noise and is determined by our cognitive abilities. Listening to speech in noise becomes much more difficult when we perform multiple tasks: when our attention becomes divided, we may not be able to devote enough conscious attention to a particular speaker and will not be able to understand the speaker as well. This is an especially challenging problem for people who are hearing impaired.

Hearing impairment (the loss of auditory sensitivity) is a significant problem for many people. Yet hearing loss can be difficult to diagnose because of conscious and unconscious adaptive strategies that people develop to understand spoken language. Individuals with mild hearing loss may not notice the loss of sensitivity because they devote more attention to auditory perception (next time you are at a party, notice how you are directing your attention when talking with someone). Many people with hearing loss even fill in missing parts of a conversation with highly predictable information based on the context of the discourse. Thus, attention (top-down) can make up for lack of auditory sensitivity (bottom-up). These adaptive strategies are remarkably successful, but will ultimately break down when the hearing loss becomes more severe. When the loss of auditory sensitivity becomes so advanced that a large amount of auditory information is unavailable, context and attention no longer provide a benefit. In such cases, a person may seek consultation with a hearing specialist and receive a diagnosis of hearing loss.

Another way in which top-down cognitive mechanisms and prior experience influence auditory perception comes from our familiarity with particular talkers. Although we may speak a common language, all talkers produce speech differently and the listener must rapidly and efficiently adapt to talkers on an individual basis. Talker familiarity can provide significant benefits when listening to speech in noise. Research has shown that individuals find talkers that they know more intelligible than unfamiliar talkers. In this case, the information that we have learned about a specific talker's voice (the talker's accent, dialect, idiosyncrasies in pronunciation, and speaking style) influences our auditory perception. In these examples, our perception is altered by what we know about a talker's voice and allows us to devote attention to those aspects that will be most useful under difficult listening conditions. Additionally, listeners have better memory for speech produced by familiar talkers compared with that of unfamiliar talkers, suggesting that familiar talkers may be afforded different status than are unfamiliar talkers. The power of such an exposure effect may be why many radio and TV commercials use voices of well-known celebrities without explicitly identifying them to promote a certain product: if the listeners are familiar with the voice of the speaker, they may pay more attention to and remember the message that the advertisers want to convey. Such indexical features (the acoustic properties of a speaker's voice that reflect his or her identity and condition, such as age, gender, size, health, dialect, etc.) can thus significantly alter our interpretation of speech.

The context under which we are listening is also particularly important for auditory perception because it will determine where we direct our attention and what "pops out" from the background. In certain environments, we would not expect to hear a certain sound and may differentially direct attention when an unexpected sound occurs. The sound of a police siren may indicate different things if you are sitting in your living room versus when it is behind you in traffic. The sound of a growling bear may mean one thing if you are at a zoo, but something completely different if you are on a hike in the woods. In both examples, the sounds are the same, but the way we perceive and interpret them and, hence, how we react to them will differ significantly because of context. A sound that may be perfectly mundane in one environment will become particularly salient when heard in a different context. Expectation can alter auditory perception in profound ways. Sine wave speech is a form of artificial speech that consists of three tones that vary in frequency, intensity, and rhythm as does the original speech signal. Yet, the perception of such stimuli depends on the instructions that the listener is given. If told that they are listening to birdsong, listeners will identify the stimulus as birdsong. However, if told that they are listening to computer-generated speech, listeners will perceive the signal as speech. In this case, the same identical stimulus can be heard in two different ways depending on the context in which we are listening. Similar top-down effects may be responsible for detecting subliminal messages in records when played backwards: If we expect to hear a secret message in “Stairway to Heaven,” then that is exactly what we perceive.

Perceptual Learning and Expertise

Perceptual learning plays a central role in auditory cognition and is especially important when learning about sound. Auditory perceptual learning is caused by exposure to and familiarity with sound and will alter our perception of sound in the future. From a young age, children are encouraged to identify sources of sound in their environment. Whether playing with toys that match a sound with a picture of the source that produced it (e.g., the sound of a cow with the picture of a cow), children are constantly refining their cognitive abilities through interactions with their environment. Understanding the sources of sounds is especially important for survival because the sounds that surround us contain information that will help to determine what actions are most appropriate. Alerting signals in the environment can indicate situations that require our attention. The sound of a smoke detector or car horn may alert us to impending danger, prompting us to react. Yet, the perception of environmental sounds is not limited to simple alerting functions. Returning to our example of the barking dog, a person who has extensive experience with dogs may interpret the sound differently than does someone who does not have much experience with dogs. Experience with sounds alters the number of dimensions that we can use to classify the sounds. A person who has grown up having dogs as pets may be able to hear many different things in the sound of a barking dog compared with a person who is unfamiliar with dogs. Through their past experience with a sound, its perceptual dimensionality (the number of different dimensions on which a sound can be classified) has increased. This is especially salient for auditory experts.

Auditory expertise comes from practice and experience. Many auditory experts perceive sound differently than do nonexperts. For example, a skilled auto mechanic may be able to listen to the sound of an idling engine and diagnose what is wrong with a vehicle. Similarly, cardiologists may be able to listen to our breathing or the beating of our hearts and determine our relative health. Parents may be able to listen to the cries of a baby and determine whether the baby is hungry, lonely, needs to be changed, or is in distress. In all of these examples, auditory expertise and experience increases the number of perceptual dimensions on which the sound can be classified. Practice listening to specific sounds will influence how we perceive them in the future, and the processes of perceptual learning are the heart of these interactions.

Another salient example of perceptual learning leading to auditory expertise comes from music. Whether listening to the radio, at a concert, or in an environment that has background music, the auditory perception of music is a part of our daily lives. Yet, nonmusicians encode and perceive music differently than do people who have had extensive musical experience. This goes beyond mere music appreciation (some people may prefer heavy metal, but others may feel that it just sounds like noise) and focuses more on the perceptual abilities of musicians (or music aficionados) and nonmusicians. A musician may be able to listen to a piece and pick out individual instruments from among a symphony. Musicians may be able to determine whether an instrument is slightly out of tune or whether the music is being played out of time. They may even be able to identify an individual musician based on the particular musical style. Some gifted musicians may even be able to listen to a piece of music once and replicate it on a particular instrument. Although many novice listeners can appreciate and enjoy music, they perceive music differently based on their previous auditory experience. Recent work in cognitive neuroscience has shown that musicians even use different brain regions when listening to music than do nonmusicians. In all of these cases, perceptual learning increases the number of dimensions on which a sound can be classified and recruits different neural systems.

Jeremy L. Loebach, Christopher M. Conway, and David B. Pisoni

See also Auditory Scene Analysis; Perceptual Learning; Speech Perception; Speech Production

Further Readings

Bregman, A. A. (1990). *Auditory scene analysis.* Cambridge: MIT Press.

Gaver, W. W. (1993). What in the world do we hear? An ecological approach to auditory source perception. *Ecological Psychology, 5*(1), 1–29.

McAdams, S., & Bigand, E. (1993). *Thinking in sound: The cognitive psychology of human audition.* Oxford, UK: Clarendon Press.

AUDITION: DISORDERS

The World Health Organisation (WHO), in its factsheet on deafness and hearing impairment, refers to hearing impairment as "the complete or partial loss of the ability to hear from one or both ears. The level of impairment can be mild, moderate, severe or profound."

The classifications are based on hearing thresholds (see Table 1, next page), most commonly measured using a standard procedure known as pure tone audiometry. In this approach, a measure of the quietest sound an individual can perceive is recorded at a number of different frequencies of sound. Plotting these values creates the audiogram.

WHO estimates from 2001 suggest that if an audiogram were to be recorded for every individual on the planet, approximately 600 million people worldwide would have a hearing loss affecting both ears, and 250 million of these would be of moderate-to-profound severity.

The actual level of disability resulting from a hearing loss is influenced by many factors beyond hearing thresholds. These factors include the amount of distortion in audible sounds, difficulties in processing and filtering auditory information, and the complexity of listening environment exposed to in a hearing impaired person's daily life. In general, a hearing loss of moderate or worse severity in both ears is deemed of particular clinical significance because, to be audible, a significant proportion of speech sounds must be raised beyond normal conversational levels. This has important consequences for the development of speech and language in children, which may be limited unless measures are taken to overcome the loss.

Additional to the level of hearing loss, the affected site or sites along the auditory pathway are used when labeling hearing impairments. Divisions are usually made into conductive (affecting the outer or middle ear), cochlear, retrocochlear (affecting the nerve of hearing), and central (affecting the cognitive processing of sound information).

This entry describes different types of hearing loss as well as some information on the relevant management options available.

Conductive Disorders

A conductive hearing loss occurs when the passage of sound between the outer ear and the cochlea is disrupted.

Obstruction of the Outer Ear Canal

Perhaps the most obvious cause of a conductive hearing loss is the inability of airborne sound waves to reach the tympanic membrane, as a result of occlusion of the ear canal. This may be caused by swelling or discharge resulting from infection (otitis externa), foreign bodies, or cerumen (earwax). Abnormalities during fetal development can result in a complete absence or a progressive narrowing of the ear canal, known as atresia or stenosis, respectively. The latter may occur later in life as a result of exostoses, bony growths commonly associated with extended exposure to cold water, such as repeated sea swimming.

Perforations

A hole, or perforation, in the tympanic membrane reduces the amount of sound energy passed from the ear canal into the middle ear bones. A

Table 1 WHO Classifications of Hearing Impairment Severity

Level of Impairment	*Hearing Loss in Decibels (dBHL)*	*Description*
Grade 0 None	25 dBHL or less	Hears whispers
Grade 1 Mild	26–40 dBHL	Able to repeat words spoken in normal voice at 1 meter distance
Grade 2 Moderate	Child 31–60 dBHL Adult 41–60dBHL	Able to repeat words spoken in raised voice at 1 meter distance
Grade 3 Severe	61–80 dBHL	Hears words shouted into better ear
Grade 4 Profound	81 dBHL or more	Cannot hear shouted voice

Source: Adapted from World Health Organisation (1991), *Report of the Informal Working Group on Prevention of Deafness and Hearing Impairment Programme Planning*. Retrieved April 16, 2009, from http://ftp.who.int/nmh/Blindness-Library/EN/Deafness/PDFdocs/91_1.pdf.

Note: The values are averages of hearing thresholds at the most significant frequencies in speech: 500 Hz, 1 kHz, 2 kHz, 4 kHz.

perforation can be caused by a shock wave traveling down the ear canal from an explosion or an impact across the surface of the ear. Alternatively, the tympanic membrane may rupture during acute middle ear infections, alleviating the intense pain associated with a "bulging" eardrum. On rare occasions, middle ear infections may progress to the air-filled spaces situated in the mastoid area behind the ear, the surgical treatment of which may result in a cavity that links the mastoid to the ear canal, often with an accompanying conductive hearing loss.

Glue Ear

The swelling of tissues within the sinuses, adenoids, and throat associated with head colds often compromises the function of the pharyngotympanic (Eustachian) tube, which is responsible for equalizing the pressure in the middle ear cavity with the atmosphere (the sensation of ears popping during a yawn or swallow). If this tube remains closed for extended periods, the cells of the middle ear will absorb gases from the cavity, reducing the air pressure in the middle ear relative to the external ear. The most common consequence is for fluid to be drawn from the cells of the middle ear cavity, filling the space with viscous mucus. This condition is known as glue ear and is extremely common in young children.

Middle Ear Growths

In rare cases, growths in the middle ear may lead to conductive hearing losses. A cholesteotoma is a growing ball of infected tissue and debris arising from the tympanic membrane and progressing into the middle ear cavity. A glomus jugulare will lead to a one-sided hearing loss and often results in a pulse being clearly heard in the affected ear because of the high number of blood vessels in the growth. Both of these require surgical removal.

Otosclerosis

Surgical management may also be considered for otosclerosis, a condition associated with reabsorption and remodeling of bone material within the structures of the ear. Commonly, the changes lead to the stapes becoming fixed in place, preventing the piston-like motion required to drive sound waves into the cochlear fluids. The resulting impairment may progress to a moderate severity.

Additional to otosclerosis are a wide variety of other possible abnormalities of the ossicles (malleus, incus, and stapes) that can result in a conductive hearing loss.

Management of Conductive Hearing Impairment

Management depends on the nature of the loss: glue ear will often resolve spontaneously, but in persistent cases, a ventilation tube (grommet) is put into the tympanic membrane. Eventually the tube will work itself free and the membrane will heal. In the case of any perforation, its size and location on the tympanic membrane will determine if it will close spontaneously or need surgical intervention using a skin-grafting procedure known as tympanoplasty.

A discontinuity along the chain of middle ear bones can be surgically reconstructed, and worn middle ear bones can be replaced by prostheses. In the case of otosclerosis, the stapes can be replaced or freed from fixation. Middle ear surgery can also be performed to insert a *middle ear implant,* a small vibrating device that can drive sound waves into the ossicles.

Hearing aids may be chosen to overcome a conductive hearing loss. However, in the presence of a chronic infection, occluding the ear canal with a hearing aid will further exacerbate the problem. In these cases, the use of a bone-anchored hearing aid (BAHA) may be considered. As the name suggests, this device is connected to the skull bones (usually by means of a titanium plate partially integrated with the bone material) and passes sound waves through the temporal bone to the cochlea, bypassing the compromised conductive pathway.

Cochlear Hearing Losses

The cochlea is responsible for transducing sound energy in the form of vibrations into neural impulses. Damage or abnormalities in the sensitive structures of the cochlea is almost always permanent, with the most common causes described here.

Congenital

Cochlear hearing loss at birth may be associated with infection (specifically cytomegalovirus and rubella), ototoxic drugs, and preeclampsia or toxemia. Genetic abnormalities can lead to hearing impairment, often as part of a syndrome.

Several countries now operate a newborn hearing screen to identify permanent hearing losses as early as possible, although in some cases deafness can be delayed in onset.

Acquired

Noise

Most cochlear hearing impairment is acquired: An example is the mild temporary hearing loss, usually accompanied by a characteristic “ringing” in the ears, after an evening of loud music. Depending on the level of exposure, the hearing should gradually recover but repeated noise exposure will lead to a permanent reduction in hearing sensitivity, affecting the mid-to-high pitch range.

Age

Presbycusis is the name given to a gradual reduction in hearing sensitivity with age. It primarily influences the same pitch range as noise-induced hearing loss and is almost universal across populations.

The permanent nature of this loss results because, unlike other cell types in the human body, inner and outer hair cells in the cochlea cannot regenerate. Researchers continue to investigate the mechanisms that result in inner and outer hair cell degeneration, with the aim of determining how these delicate structures can be replaced and the damage to them prevented, slowed, or repaired.

Sudden Onset

In cases of sudden-onset hearing losses, emergency attention from an ear, nose, and throat (ENT) specialist is required within 48 hours of occurrence. If there is no obvious cause such as traumatic injury or severe infection (e.g., bacterial meningitis), then investigations will be undertaken to identify possible causes, which may be viral, vascular, or autoimmune (the latter referring to an erroneous immune-mediated attack of the structures of the inner ear). Steroids or antivirals may be prescribed to limit permanent damage to the inner ear, and in many cases, the hearing may fully or partially recover, even in cases where the exact cause remains unknown (idiopathic).

Perilymph Fistula

A small hole may develop in the membranes of the inner ear as a result of trauma from middle ear surgery, a sudden change in pressure (such as during scuba diving), or intense physical strain. If this leads to the escape of cochlear fluid, it is known as a perilymph fistula. The most common sites for this to occur are the oval and round windows, leading to compromised hearing and often accompanied by tinnitus and dizziness. The symptoms may be exacerbated by coughing, sneezing, changes in middle ear pressure, and possibly loud noises (Tullio phenomenon).

Ménière's Disease

Another disorder thought to be associated with abnormal pressure of the internal fluids of the cochlea is Ménière's disease. For reasons yet to be fully determined, researchers believe that there are spontaneous periods of abnormally high pressure within the cochlea that distort the delicate internal membranes. This results in the simultaneous occurrence of three or more of the following symptoms: hearing loss, tinnitus, an intense feeling of spinning, and a pressure or fullness in the ear.

Ototoxicity

A number of drugs have a toxic effect on structures in the auditory system, especially the cochlea. These include some antibiotic and chemotherapeutic treatments, and as a result hearing loss may be an unavoidable consequence of the treatment of life-threatening diseases. The extent and nature of ototoxic hearing loss is influenced by a wide variety of factors, specific to both the individual and the treatment regime, and research continues to further understand the biochemical and physiological processes involved in drug ototoxicity, an in-depth discussion of which is beyond the scope of this entry.

Management

Dead Regions and Profound Hearing Loss

Loss of inner hair cells in a particular region of the basilar membrane will result in an inability to encode sound at that particular pitch, known as a cochlear dead region. If inner hair cell damage is widespread, there may be no significant benefit from a standard acoustic hearing aid. If this is the situation in both ears, nonverbal methods of communication such as sign language and lip-reading may be adopted, or a cochlear implant may be considered. Alternatively, a similar device can be implanted into the auditory brainstem if the nerve connections to the cochlea are compromised.

Mild-Moderate Hearing Losses

Mild-moderate cochlear hearing loss is associated with damage to outer hair cells, which are responsible for preferentially amplifying low-level sounds. This leads to the inaudibility of quieter components of speech, although often sounds of greater intensities remain about as loud as they would be heard in the absence of a hearing loss, and in some cases are actually perceived as being louder. This effect is known as recruitment.

A number of features on digital hearing aids aim to overcome the complexities of cochlear hearing loss. Even so, some hearing aid users with cochlear hearing loss remain unable to hear speech clearly, and most will still have difficulty understanding speech in noisy environments. This is partly because of a compromise of other auditory abilities including the ability to discriminate between different pitches of sound, and the ability to follow changes in sound level over short time intervals.

Retrocochlear

The retrocochlear portion of the auditory pathway refers to the nerve that carries auditory information (the VIIIth cranial nerve and the point at which it connects to the brain). Retrocochlear lesions have several possible causes, including the following:

- Immaturity caused by premature birth
- Degeneration caused by multiple sclerosis
- Vestibular schwannoma

The last example is a benign tumor that develops from the cells that insulate the vestibular portion of the VIIIth cranial nerve. To avoid life-threatening consequences, the growth may be surgically removed, which (depending on the technique used) may result in a complete loss of hearing on the affected side.

Central

Abnormalities in the cognitive processing of auditory information may result in difficulties with speech comprehension or sound localization, even though hearing thresholds are "normal." The origin of the problem may be damage to the auditory cortex caused by trauma or a stroke.

In cases where there is no obvious cause for abnormally poor auditory performance, the test batteries that exist to quantify these difficulties are described by the collective term *auditory processing disorders* (APD). Those suffering from APD often find speech comprehension to be improved when the impact of background noise is limited and visual information, such as lip-reading, is used to supplement auditory information.

When a hearing loss is present without intervention for an extended time, the resulting deprivation of stimulus leads to a reorganization in the central auditory system (this ability to organize and reorganize in response to long-term changes in sensory input is known as plasticity). This might limit the performance of a hearing aid or implant tried later. Additionally, auditory plasticity in the early years of life facilitates the ability to perceive and process spoken language, and this is one of the reasons why newborn hearing screening programs aim to identify hearing impairments as early as possible, at a stage where intervention can make use of this critical period to maximize the resulting auditory ability.

Phil Gomersall and David Baguley

See also Audiology; Audition; Audition: Loudness; Audition: Pitch Perception; Auditory Processing: Central; Auditory Processing: Peripheral; Auditory System: Damage Due to Overstimulation; Auditory System: Structure; Auditory Thresholds; Cochlear Implants: Controversy; Cochlear Implants: Technology; Hearing Aids; Tinnitus

Further Readings

Ballantyne, J. C., Graham, J. M., & Baguley, D. M. (2009). *Ballantyne's deafness* (7th ed.). Hoboken, NJ: Wiley.

Corbeel, L. (2007). What is new in otitis media? *European Journal of Pediatrics, 166,* 511–519.

Luxon, L. (Ed.). (2003). *Textbook of audiological medicine.* London: Martin Dunitz.

Tait, M. E., Nikolopoulos, T. P., & Lutman, M. E. (2007). Age at implantation and development of vocal and auditory preverbal skills in implanted deaf children. *International Journal of Pediatric Otorhinolaryngology, 71,* 603–610.

World Health Organisation (1991). Report of the Informal Working Group on Prevention of Deafness and Hearing Impairment Programme Planning. Retrieved from http://www.who.int/entity/en

World Health Organisation (2006). Factsheet No. 300: Deafness and Hearing Impairment. Retrieved from http://www.who.int/mediacentre/factsheets/en

AUDITION: LOUDNESS

Loudness is a subjective attribute of sounds that allows us to experience the dynamics of music, speech, and our environment. Sounds can range from soft and calming to so loud that they startle us. The louder the sound, the faster we respond to it. Loud sounds draw our attention, alert us to danger, and pump us up when we are working out at the gym.

Exposure to very loud sounds can also cause hearing loss. When most sounds reach potentially damaging levels, they become unpleasant and we refrain from exposing ourselves to them. The exception to this rule is music. The normal controls that protect us from sound exposure that can result in hearing loss seem to be absent for music, but only if we find the type of music pleasant. Loud music can be mood altering and is often sought after. This appears to be true across a number of different cultures and may be universal.

Problems arise not only when a sound is too loud, but also when a sound is too soft. We may not be aware of a soft sound that we want or need to hear because it blends in with background sounds. Attending to a soft sound requires effort, and we strain to hear. Listening to a sound at a comfortable loudness is preferred.

Achieving a comfortable loudness requires an understanding of the many factors that influence the perception of loudness. Loudness changes with the physical properties of sound and the physical and psychological states of the perceiver. The influences of these properties are described in this entry following a definition of loudness and a short description of the methods used to measure loudness. This entry also includes a description of our ability to detect and discriminate small changes in level and concludes with an overview of models of loudness.

Language of Loudness: Definitions and Terminology

Although we know what loudness is, we may not be aware of the exact definition and terminology used for loudness. Loudness is the perceptual strength of a sound that ranges from very soft (or quiet) to very loud. Loudness is the *primary* perceptual correlate of the level of a sound. The study of loudness is a subarea of psychoacoustics. *Psychoacoustics* is the study of the relationship between physical properties of sound (e.g., frequency, level, and physical duration) and our perceptual response to them (e.g., pitch, loudness, and subjective duration).

Loudness is often erroneously described as the "volume" or "intensity" of a sound. The term *volume* is used to describe a percept different from loudness and should not be confused with loudness. Volume refers to the subjective size of a sound. The term *intensity* refers to a physical property of a sound, which is related to its level. A common unit of measure is sound pressure level in decibels, or dB SPL, which is the logarithmic ratio of intensity. Therefore, the terms *volume* and *intensity* should not be used interchangeably with *loudness.*

Since the late 18th century, it has been common for composers to use written notation to indicate the level at which a piece of music should be played to produce a corresponding loudness. These are called dynamic markings and are usually in Italian. Examples of dynamic markings are *pianissimo* (very soft), *mezzo piano* (medium soft), *forte* (loud), and *fortissimo* (very loud). These dynamic markings refer to relative levels, rather than to specific levels, so that a passage of music may be a little louder or softer than another. Music notation also includes words to indicate a general increase (*crescendo*) or decrease (*decrescendo* or *diminuendo*) in loudness. In addition, a composer can indicate that one note should be louder than the other notes by using accents. Such notations help to give music the proper loudness dynamics and indicate the beat of the music.

Measuring Loudness and Units of Loudness

Because loudness is a subjective attribute, it cannot be measured directly. Two common methods used to indirectly measure loudness are equal-loudness matching and scaling. In equal-loudness matching, a physical property of one sound is adjusted so that the sound is judged equal in loudness to a fixed reference sound; usually the two sounds are presented alternately. Equal-loudness matches were used to derive the phon scale. The phon is a unit of loudness level that indicates the level of a tone required for it to sound equal in loudness to a 1,000-hertz (Hz) reference tone. For example, the loudness of a 70-phon tone at any frequency is equal to the loudness of a 1,000-Hz tone at 70 dB SPL.

In scaling methods, a sound is presented to a person many times at different levels that encompass a wide range of levels. After the presentation of each sound, the person's task is to provide information about the loudness elicited by the sound. One common task is to ask a person to assign a number to the loudness of a sound. This number is called a *loudness estimate* and the method is called *magnitude estimation.* The sone scale was determined by magnitude estimation. One sone is the loudness corresponding to a 1,000-Hz tone at 40 dB SPL, which is the same as a loudness level of 40 phons. In another task used to obtain information on loudness, a person is asked to adjust a physical property of a sound to produce a certain subjective magnitude of a sound. This is called *magnitude production.* The final estimate is an average of data obtained with both methods. In another type of scaling method, cross-modality matching, a person is asked to match the loudness of a sound to a perceived property in another modality, such as brightness or perceived line length. For example, a person may be asked to cut the length of a string as long as a sound is loud. In still another method, five to seven fixed category labels are used (such as "very soft," "soft," "comfortably loud," "loud," "very loud," and "too loud") and a person is asked to assign the loudness of a sound to a preselected category. This method is called *category scaling.*

Loudness and the Physical Properties of Sound

Loudness changes as a function of level, frequency, bandwidth, duration, spectral complexity of a sound, and the presence of other sounds. There is no simple one-to-one correspondence between loudness and any physical property of a sound, including level. Audible sound levels that do not immediately damage hearing range from roughly 0 to 120 dB SPL. In general, loudness will double for every 6 to 10 dB increase in SPL, but does not increase indefinitely. As the level is increased beyond being "too loud," sensation transitions from a "tickle" feeling to pain. Research on loudness at such high levels is no longer performed because of ethical concerns.

The way in which loudness increases with level has been measured using magnitude estimation. A plot of loudness estimates on a logarithmic scale as a function of audible sound levels in dB SPL is known as a *loudness function.* The loudness function has been described as $L = kI^e$, where L = Loudness, k = a constant, I = Intensity, and ^{e} = the exponent, which is proportional to the slope of the function. Although simple, it is not the most accurate description. There are two important deviations. The first is that loudness increases rapidly with increasing level near threshold. The second is that loudness changes more slowly with increasing level at moderate levels than at high and low levels. These two deviations result in an entirely different psychophysical law that relates level to loudness (i.e., a nonstationary point of inflection law or an inflected exponential [INEX] law).

Loudness changes throughout the audible frequency range from about 20 to 20,000 Hz. A plot that shows the levels required to keep loudness constant as frequency changes is called an equal-loudness level contour (or equal-loudness contour). Most textbooks and the recommended readings at the end of this entry present a family of these contours that range from near threshold (about 10 phons) to high levels (about 100 phons), in steps of 10 or 20 dB. Near threshold, the equal-loudness level contours follow absolute threshold. This means that the levels of tones at very low and very high frequencies need to be set to higher levels than the 1,000-Hz reference tone to sound equal in loudness to it. For example, a 125-Hz tone at 34 dB SPL is equal in loudness to a 1,000-Hz tone at 10 dB SPL, and both have a loudness of 10 phons. Therefore, a 24-dB difference in level is required for the two tones to be perceived as equally loud. At high levels, the equal-loudness level contours flatten, which means that the difference required to set the two tones to equal loudness is much less. For both tones to have a loudness of 100 phons, the 125-Hz tone needs to be about 104 dB SPL and the 1,000-Hz tone needs to be only 4-dB lower for the two tones to be perceived as equally loud.

Loudness increases with bandwidth (frequency range) and duration of sounds, but only within certain physical parameters. If the overall level of two tones is kept constant and the frequency separation between them is increased, loudness of the two tones heard together will remain relatively unchanged until the frequency difference between the tones exceeds about 20% of the geometric average of the frequency of the two tones, after which loudness increases. This is called spectral loudness summation. It occurs for sounds as the bandwidth is increased. The magnitude of increase changes with level and is greatest at moderate levels, so that a 1,000-Hz tone would need to be as much as 18 dB SPL higher than a sound encompassing a broad range of frequencies for the two sounds to be perceived as equally loud. Loudness also increases as the duration of a constant-level sound increases up to about half a second, after which loudness remains relatively constant. This is called temporal integration (or temporal summation) of loudness. The magnitude of this effect can be as much as 15 to 20 phons.

Loudness changes with the spectral complexity of a sound and the presence of other sounds. The details of the influence of spectral complexity is beyond the scope of this entry, but sounds that come before and after a sound can affect its loudness. Finally, loudness changes depending on the background against which a sound is heard, or the background "soundscape." For example, a friend's voice will sound softer on the street near a noisy construction site than in a quieter area. This phenomenon is referred to as partial masking of loudness because the loudness of a sound that you are listening to (i.e., a friend's voice) is partially masked, or blocked, by the background sounds.

Physiological and Psychological States of the Perceiver

Loudness changes as the physical and psychological states of the perceiver change. For example, loudness can change depending on whether a person listens with one or two ears, whether the ears have been exposed to loud sounds, whether a hearing loss is present, and whether a person is paying attention to the sound. A person with normal hearing in one ear and a hearing loss in the other ear will perceive sounds as less loud than when listening with two normal ears. This occurs because loudness summates across the two ears. This phenomenon is called *binaural loudness summation.* Although sounds heard by two ears are louder than sounds heard by only one ear, the amount that loudness increases is still a matter of debate and appears to change with the type of sound and listening environment.

Loudness, as well as threshold, may also change after a person is exposed to high-level sounds. This change in loudness can be temporary, taking anywhere from a few minutes to several days to return to normal. For example, sounds may be perceived softer than normal after attending a rock concert without hearing protection, but loudness returns to normal by the next day. This temporary condition is called *fatigue,* or *temporary loudness shift.* However, with exposure to very loud sounds over a prolonged time, or even a short time with sufficient intensity, changes in loudness can be permanent.

Loud sounds draw our attention and we instinctively turn our attention toward them. Attention plays a role in our perception of loudness. For example, evidence indicates that sounds that are not attended to may be perceived as less loud than sounds that have our full attention. Furthermore, a sound that is continuously presented at a low level for a long period will fade into inaudibility. This is called *loudness adaptation.*

Detection and Discrimination of Level Differences

This section describes the impressive ability of the auditory system to detect changes in level (level detection) and discriminate small level differences (level discrimination or intensity discrimination). The topics of level detection and level discrimination are usually raised in conjunction with discussions of loudness. This is a debatable practice because the subjective attributes of changes in small level differences may not be perceived as differences in loudness, they may be perceived as changes in pitch or other subjective attributes of sound. Level detection and discrimination are included in this entry because of the implications of these types of experiments for understanding the physiological encoding of level information in the auditory system.

The smallest audible difference in level between two sounds has been called the just noticeable difference (JND) or the difference limen (DL). The JND has been measured by increasing and decreasing the level difference until it is just audible using three types of tasks: (1) discrimination of two sounds separated in time that differ only in level, (2) detection of amplitude modulation (i.e., variations in strength) of a sound, and (3) detection of an increment in the level of a continuous sound. Although the exact values of the JND differ somewhat among the different tasks, the general trends in the data are the same.

The JND for level changes as a function of level, frequency, and duration of the sound. One of the oldest scientific observations in hearing describes how the JND for intensity changes with increasing intensity of a noise. It is called *Weber's law* (also known as Weber's fraction) and refers to the fact that the JND is proportional to the physical magnitude of the sound. In other words, the just-noticeable change in intensity divided by the intensity of the stimulus remains constant with increasing level. This means that the JND is about 0.5 to 1 dB for all levels, except the lowest levels. Weber's law does not work at low levels for noises, or as well for tones as it does for noises. For tones, there is a deviation from this proportional relationship, called the *near miss to Weber's law.* This deviation is a decrease in the Weber fraction with level, which results in a smaller JND in decibels at high levels.

An excitation pattern model provides a good theoretical framework for describing our ability to discriminate small level differences. According to this model, any audible sound creates a pattern of activity in the auditory system. These activity patterns are called excitation patterns. As the level of a tone is increased, its excitation pattern increases and spreads to adjacent frequencies. When a critical difference is reached between the excitation patterns of two tones that differ in level, the auditory system has sufficient information to discriminate the sounds. For example, a low-level tone has a small excitation pattern and requires a 2-dB level difference between two tones to create a difference between the two excitation patterns large enough to reach the threshold of discrimination. A high-level tone has a large excitation pattern that spreads to adjacent frequencies and requires only a 0.5-dB level difference between the two tones to create a difference between the two excitation patterns large enough to reach the threshold of discrimination.

The theoretical framework of excitation patterns can also explain the effect of frequency on level discrimination. The JND for level is larger for frequencies close to the high-frequency limit of audible sounds. According to the model, this occurs because the excitation patterns are effectively cut off because of the frequency limits of audibility. In the absence of the spread of excitation at the high frequencies, the JND must be increased to create a difference between the two excitation patterns large enough to reach the threshold of discrimination.

If you recall temporal integration (or temporal summation) of loudness described earlier, it is not surprising that the auditory system sums up information over time. This summation also occurs for level discrimination. As the duration of a tone increases up to about one-half to one second, the level difference required for discriminating two tones decreases.

Although the basis for coding of level in the auditory system is not completely understood, any model of coding must consider the impressive ability of the auditory system to detect and discriminate relatively small level differences. Much of what is known about the physiology of level perception comes from correlating psychoacoustic measurements with the physiological responses to level differences. For example, information from a single neuron in the auditory nerve is enough to account for our ability to discriminate two sounds that differ in level. However, just because information is available does not mean that it is used by the auditory system. Exactly how information is carried from the auditory nerve to the brain is not well understood, but the models described in the next section may give us some insight.

Models of Loudness

Although no model of loudness can account for all the phenomena associated with the perception of loudness at the time of this entry, models exist that can explain certain aspects of the perception of loudness. These models can be divided into two types: models that describe and predict the relationship between the stimulus and the perception of loudness (i.e., psychoacoustic models) and models that attempt to make correlations between changes in the level of a stimulus and the physiological response to these changes (i.e., physiological models). These two types of models are described here.

As described earlier, an inflected exponential function is the most accurate descriptive psychoacoustic model of how the loudness of a tone increases with increasing level. However, most sounds are much more complex than a tone. Complex models have been developed to predict the loudness of sounds from the physical spectrum. Basically, these models include three stages: the physical spectrum is divided into frequency bands, a weight is assigned to each frequency band, and the weighted outputs of the individual bands are summed. The most recent models also consider filtering of the signal by the ear and listening condition. Loudness models were initially developed for steady-state sounds and people with normal hearing, but variations of these models exist to account for the loudness of time-varying sounds and the presence of hearing loss. These models are still under development and do not presently account for all conditions.

Physiological models of loudness are less developed than psychoacoustic models are. Only some aspects of the physiology of loudness appear to be explained. For example, the increase in loudness with increasing level is consistent with the basilar-membrane response function. There is a good correlation between the inflected exponential function and physiological data. However, how the loudness of a sound increases with level is less well understood at the auditory nerve. As the level of a tone increases, the firing rate of neurons in the auditory nerve also increases, but most neurons tend to saturate within a dynamic range of only 30 to 60 dB so that increases in level cause no further increase in the firing rate. Although some benefit is obtained from a small number of neurons with higher thresholds, this does not appear to be enough to account for the fact that loudness increases over a level range of about 120 dB SPL. This is known as the dynamic range problem.

Although our understanding of the physiology of loudness is far from complete, two compelling lines of thought that are not mutually exclusive explain the dynamic range problem. The first is that loudness is related to the total amount of neural activity. The second is that loudness is related to the temporal properties of the neural activity. As a tone increases in level, it excites neurons with primary sensitivity (characteristic frequency or the frequency to which it responds best) near the tone frequency, and it excites an increasing number of neurons with adjacent characteristic frequencies. This is known as the *spread of excitation.* Temporal theories are based on the fact that neurons that respond to a certain frequency tend to phase lock to it. In other words, they produce neural firings at precise times correlated with temporal properties of the sound wave. When the level of a tone increases, more neurons phase lock to it and the overall synchrony across the population of auditory nerve fibers increases. However, the timing theories have difficulty at high frequencies because the ability of the auditory nerve fibers to phase lock decreases at high frequencies, which is known as the roll-off in phase locking. It is possible that the amount and range of neural activity, as well as the timing and phase locking of neural activity, play roles in the perception of loudness. Despite how much data have been obtained, surprisingly little has been proven regarding the connection between physiological responses and our perception of loudness.

There is a need for models that can predict the average perception of loudness for large groups of listeners, such as those in concert halls and other public spaces. A sound-level meter provides information about the level of a sound, but does not provide information about loudness. Attempts have been made to design "loudness meters" that provide a number corresponding to the subjective experience of loudness for an average group of people with normal hearing. Because loudness can vary among individual listeners, measures from these meters do not apply to any one individual person. Whereas meters may provide a good estimate of the group average for steady-state sounds, scientists are still in the process of designing meters for time-varying sounds that constitute most environmental sounds.

At the time of this entry, there is no comprehensive theory that can explain all phenomena related to our ability to perceive loudness—this amazing ability that allows us to experience the dynamics of music, speech, and sounds in our environment.

Mary Florentine and Michael Heinz

See also Attention: Effect on Perception; Audiology; Audition; Audition: Cognitive Influences; Individual Differences in Perception

Further Readings

Epstein, M., & Marozeau, J. (in press). Loudness and intensity coding. In C. Plack (Ed.), *Oxford University Press handbook of auditory science: Auditory perception.* Oxford, UK: Oxford University Press.

Fastl, H., & Zwicker, E. (2007). *Psychoacoustics: Facts and models* (3rd ed., Chapter 8). New York: Springer.

Florentine, M., Fay, R., & Popper, A. (Eds.). (2009). *Loudness.* New York: Springer.

Green, D .M. (1993). Auditory intensity discrimination. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Human psychophysics* (pp. 13–55). New York: Springer.

Heinz, M. G., Colburn, H. S., & Carney, L. H. (2001). Rate and timing cues associated with the cochlear amplifier: Level discrimination based on monaural across-frequency coincidence detection. *Journal of the Acoustical Society of America, 110,* 2065–2084.

Moore, B. C. J. (2003). *An introduction to the psychology of hearing* (5th ed., Chapter 4). Boston: Academic Press.

Plack, C. J., & Carlyon, R. P. (1995). Loudness perception and intensity coding. In B. C. J. Moore (Ed.), *Hearing* (pp. 123–160). San Diego, CA: Academic Press.

Relkin, E. M., & Doucet, J. R. (1997). Is loudness simply proportional to the auditory nerve spike count?

Journal of the Acoustical Society of America, 101, 2735–2740.

Scharf, B. (1997). Loudness. In M. J. Crocker (Ed.), *Encyclopedia of acoustics* (Vol. 3, pp. 1481–1495). Wiley.

AUDITION: PITCH PERCEPTION

Pitch is a subjective attribute of sound that allows one to enjoy a musical melody, differentiate between talkers, and identify many different sound sources. Sound itself does not have a pitch. However, without an ability to perceive pitch, speech would be Morse code (only changes in the loudness or duration of a sound could be used to convey language), music would be only drum beats, and the sound from most sources would go unrecognized. This entry examines the physical dimensions of sound, discusses how pitch is measured, and describes some of the relationships between the physical dimensions of sound and perceived pitch.

Sound

Sound has the physical dimensions of frequency, level, and time. Sound is caused by objects that vibrate and send out sound waves. *Frequency* refers to the rate at which an object vibrates, and *level* refers to the magnitude of that vibration. Hertz (Hz) is the physical measure of frequency. A sound with a 100-Hz frequency is a sound whose vibratory pattern repeats 100 times per second; if the vibration occurs at rate of 1,000 times per second, the sound's frequency is 1,000 Hz, and so on. The perception of pitch is the auditory system's means of processing information about the physical properties of sound: frequency, level, and time. When the physical frequency of the sound from a source varies, it is almost always perceived as a pitch change. A short string on a violin vibrates faster than a longer string does and produces a high-frequency sound that is perceived as a high pitch. But many other aspects of frequency, level, and time can lead to the perception of pitch. The study of pitch perception involves understanding the relationship between the physical properties of sound and the subjective perception of pitch. As of yet, there is not a complete understanding of these relationships and, as a consequence, the study of pitch perception is an active area of research in the auditory sciences.

Measuring Pitch

The standard definition of pitch is that pitch is that subjective attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high. Pitch is measured subjectively in three ways: using musical scales, using a frequency scale, and occasionally using the mel scale. The musical scale contains notes (in Western music, 12 notes: A, B, C, D, E, F, G, and in the sharps, #, and flats, b: Ab, Bb, F#, A#, Eb). The musical scale may be divided into octaves. An octave is a doubling of frequency (880 Hz is the octave above 440 Hz, and 220 Hz is the octave below 440 Hz; 440 Hz is the approximate vibratory frequency of the musical note A played in the middle octave of the piano keyboard). Each octave can be divided into 12 equal logarithmic divisions called semitones. Thus, all the A notes on a piano keyboard are separated from each other by an octave, and each of the 12 notes (e.g., black and white keys on the piano keyboard) within an octave are spaced in frequency by about a semitone. Each semitone has 100 cents (hence an octave has 1,200 cents). Musical notes, semitones, and cents are used as scales to measure pitch.

In many experiments, the pitch of a sound with a known pitch (the comparison sound) is matched to that of a sound with an unknown or to-be-measured pitch (the target sound). In such pitch-matching procedures, listeners are asked to adjust a physical property of the comparison sound (e.g., frequency) so that the comparison and target sound have the same perceived pitch. Usually a pure tone (a periodic sound) or a periodic series of clicks (click train) are used as comparison sounds. The frequency of the pure tone or the repetition frequency of the click train, both expressed in Hz, are used as the scale or measure of pitch. That is, Hz is a measure of pitch, just as it is a measure of frequency. However, physical frequency and subjective pitch are not the same thing, as several examples will indicate.

In some cases, the mel scale is used to measure pitch. The mel scale is derived from a subjective scaling procedure such that the pitch of a tone presented with a frequency of 1,000 Hz is defined as having a pitch of 1,000 mels, and the 1,000-mel sound is the standard pitch sound. If a sound is perceived to have a pitch that is half that of the pitch of the standard, then the pitch is 500 mels. A sound with twice the pitch of the standard would have a 2,000-mel pitch.

Different Meanings of Musical Pitch

Musical pitch can take on two different meanings. If one starts at the left most piano key on the keyboard and one strikes each key in succession moving from left to right, each key will produce a higher pitch than the preceding note. In addition, the pitch of each piano key in one octave has the same pitch as the corresponding key in another octave (e.g., each A key across the key board has the pitch of A), in the sense that a recognizable melody can be played in any octave (the basis of musical transposition). Thus, the A key to the left of the piano keyboard has both the same pitch and a lower pitch than that of the corresponding A key to the right on the keyboard. *Pitch height* is used to refer to perceived musical pitch that occurs independent of the octave (e.g., each key on the piano keyboard produces a different pitch height), and *pitch chroma* refers to the perceived musical pitch of corresponding notes across octaves (e.g., all of the A keys on the piano keyboard have the same pitch chroma).

Humans, especially those with musical training, and several animals appear to perceive octave relationships among pitches as unique from non-octave arrangements. The pitch of a sound can support melody recognition; such pitches are referred to as *musical pitches* and are considered the strongest form of pitch perception. Other sounds produce pitches that are poor for melody recognition but listeners can accurately judge musical intervals (semitones), and other stimuli do not support either melody or musical interval recognition, but the pitch of a pure tone can be matched to the stimulus. In some cases, the pitch of a sound is weak leading mainly to a change in the perceived quality (timbre) of the sound.

Listeners who can reliably label (or produce by singing) the correct musical note of a sound without a referent sound are said to have an *absolute pitch* ability. Some listeners who do not have absolute pitch possess a relative pitch ability in that they can reliably judge the musical interval (semitone) between two sounds. And many listeners have neither an absolute nor a relative pitch judging ability. Whether the causes of these different pitch abilities are genetic factors or experience and training is not clearly understood.

Perceived Pitch

The pitch of a simple sound is directly related to the sound's physical frequency. For instance, the pitch of a pure tone is equal to the tone's frequency. The pitch of most periodically repeating sounds is equal to the reciprocal of the period of the repetition, expressed as hertz. For instance, a series of brief sounds (transients) that are periodically repeated produce a pitch related to the reciprocal of the period of the sound. If this click train of transients has a periodic repetition of 100 Hz, then the period between each transient in the click train is .01 seconds (i.e., 1/100 Hz = .01 seconds or 1/.01 seconds = 100 Hz). Because the period of most audible sounds is short (fractions of a second), period is often expressed as milliseconds (msec), where 1 millisecond is 1/1,000 of a second. Thus, in the previous example, a period of .01 seconds is also 10 msec (1,000 × .01 = 10 msec).

The pitch of sound can also change as a function of sound level. In the physical description of sound, a sound wave's frequency is independent of its level. Level is often correlated with the subjective attribute of loudness. The subjective attributes of pitch and loudness can change with changes in either a sound's frequency or level. So although frequency and level are physically independent, either can lead to a change in perceived pitch. Composers have been known to write music with an increase in the level (loudness) of a note to also produce a slight increase in the perceived pitch of that note. Thus, even though physical level is independent of physical frequency, subjective pitch can be altered by changes in either physical frequency or level. This makes the relationship between the physical dimensions of sound (frequency, level, and time) and perceived pitch complicated.

Any complex sound can be expressed as a sum of simple or tonal sounds, where each tonal component (spectral component) has a physical frequency.

A click train is a complex sound that can be described as the sum of many different spectral components. A click train will have a fundamental vibratory frequency that is equal to the reciprocal of the period between the transients in the click trains. And, the click train is made up of the fundamental spectral component and all of its higher harmonics. A harmonic is an integer multiple of the fundamental frequency. Thus, a click train with transients that repeat once every 10 msec (once every .01 seconds) has a fundamental frequency of 100 Hz (1/.01 seconds = 100 Hz) and all of the higher harmonics of 100 Hz (200 Hz, 300 Hz, 400 Hz, 500 Hz, etc.). All musical instruments, including the human voice, produce such harmonic complexes, and the pitch of a note played by any instrument is equal to the frequency of the fundamental. Thus, the A note in the middle octave of the piano keyboard produces a harmonic complex with a 440-Hz fundamental frequency and many of the higher harmonics of 440 Hz (e.g., 880 Hz, 1,320 Hz, 1,760 Hz). The perceived pitch of this A note is 440 Hz. Many other sound sources (e.g., a rotating motor) also produce harmonic complexes and if the sound source produces a perceived pitch it is usually associated with the fundamental frequency of the harmonic complex. For example, a motor may rotate 120 times per second, which produces a fundamental frequency of 120 Hz and many of the higher harmonics of 120 Hz. In many cases, the whine of the motor would be perceived as having a 120-Hz pitch.

Thus, the pitch of such harmonic complexes could be explained by several properties of the sound. The frequency of the fundamental of the harmonic complex could be used to account for pitch. Or, the repetition rate of the sound's overall vibratory pattern could be used to account for pitch. Or, that the spacing between each successive harmonic of the sound is equal to the fundamental frequency might be a basis for the pitch of a harmonic complex (e.g., a harmonic sound with 100-Hz fundamental will have harmonics at 200 Hz, 300 Hz, 400 Hz, 500 Hz, etc., each spaced 100-Hz apart).

An interesting experiment conducted in the 19th century indicates that the pitch of harmonic complexes probably does not result from the fundamental frequency itself. A particular type of sound source (an optical siren) produces a sound with the harmonics of a fundamental, but the fundamental is not physically present in the sound's spectrum. For example, the optical siren could produce a sound with spectral components of 200 Hz, 300 Hz, 400 Hz, 500 Hz, but not 100 Hz (the fundamental frequency). The fundamental frequency (100 Hz) is missing from this sound's spectrum. The perceived pitch of the siren is 100 Hz, even though there is no physical spectral component at the fundamental. This is an example of the missing fundamental pitch. Many musical instruments (including the human voice) produce a missing fundamental pitch in that the spectral component at the fundamental is either not present or weak compared with the higher harmonics in the sound's spectrum produced by the instrument.

Accounts and Theories of Pitch Perception

Thus, the frequency of the fundamental cannot be used to account for perceived pitch, because stimuli with no fundamental in the sound's spectrum still produce a pitch equal to the missing fundamental. Harmonic complexes without the fundamental still vibrate with a period that is equal to the reciprocal of the missing fundamental (e.g., a sound containing spectral components of 200 Hz, 300 Hz, 400 Hz, 500 Hz, etc., will vibrate with a period of 10 msec or .01 seconds, the reciprocal of 100 Hz, which is the frequency of the missing fundamental). And the spacing between the spectral components of this 100-Hz missing fundamental harmonic complex is still 100 Hz. Thus, either the repetition period or the spectral spacing of the harmonics could be used as a basis to account for the pitch of the missing fundamental.

Using aspects of the temporal repetition of the vibrating sound to account for pitch perception is a temporal explanation of pitch, and using aspects of the spectral spacing of the harmonics is a spectral or frequency explanation of pitch processing. The history of attempts to develop theories and models of pitch perception can be described as a competition between spectral and temporal explanations. Some hearing scientists have also suggested that some combination of spectral and temporal approaches might provide a coherent theory of pitch perception. Currently, no one theory has been universally accepted as accounting for pitch perception.

The auditory nervous system preserves information about both the temporal repetition of sounds and the spectral content of sounds. That is, auditory nerve fibers produce neural discharges that repeat at a rate that matches the vibratory repetition of the incoming sound's waveform (to repetition rates as high as 5,000 times per second). And the auditory nervous system preserves information about the harmonics of a complex sound (at least for low-frequency harmonics). Thus, current models and theories of pitch processing attempt to account for the perceived pitch of sound based on what aspects of the physical properties of the sound (e.g., repetition period or spectral harmonics) are preserved and processed by the auditory nervous system. Several competing models of pitch perception are based on such neural processing assumptions. Thus, at present no single theory of pitch perception can account for all of the data that have been collected regarding what is arguably the most important perceptual aspect of sound, pitch.

William A. Yost

See also Absolute Pitch; Audition: Pitch Perception; Auditory Frequency Selectivity; Music Cognition and Perception

Further Readings

Cohen, M. A., Grossberg, S., & Wyse, L. L. (1995). A spectral network model of pitch perception. *Journal of the Acoustical Society of America, 98,* 862–879.

de Boer, E. (1976). On the "residue" and auditory pitch perception. In W. D. Keidel & W. D. Neff (Eds.), *Handbook of sensory physiology* (pp. 479–583). New York: Springer.

Meddis, R., & Hewitt, M. J. (1991). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. *Journal of the Acoustical Society of America, 89,* 2866–2882.

Plack, C., Oxenham, A., Fay, R., & Popper, A. (Eds.). (2005). *Pitch: Neural coding and perception.* New York: Springer.

Plomp, R. (1967). Pitch of complex tones. *Journal of the Acoustical Society of America, 41,* 1526–1533.

AUDITION: TEMPORAL FACTORS

Most sounds are temporally complex, marked by variations over time in acoustic characteristics or events. Both speech and music represent common examples of sounds that vary over time with change of characteristic and of content. Extraction of the information contained in temporally complex sounds requires response by the auditory system to these changes. Resolution, the ability to temporally resolve or follow stimulus variations, is a central aspect of the study of auditory temporal factors. Often, however, auditory perception is governed by a summation or accumulation of information over time. Thus, a second temporal factor is integration. Finally, a third aspect that incorporates both resolution and integration is the study of the perception of a sequence of resolved acoustic events in terms of pattern processing. This entry describes temporal factors of audition such as time, frequency, and modulation; resolution; integration; and auditory segregation and pattern processing.

Time, Frequency, and Modulation

Sound is a pressure waveform that is generated by a vibrating object. With sound a waveform that varies over time, temporal processing by the auditory system is a central basis for determining the information conveyed by sound. The vibrations that generate most sounds repeat in some manner or another. These periodicities are described in terms of repetition frequency. Repetitions per second are defined in units of hertz (Hz). For example, a simple or sinusoidal vibration that repeats one thousand times a second has a frequency of 1,000 Hz.

Representation of the periodicities in the time waveform in terms of frequency illustrates the relationship between the temporal and frequency-based descriptions of sound. The implication of this relationship, formalized by the Fourier theorem, is that change in one domain affects the other. The auditory system codes both the temporal and frequency characteristics of sound. Time is coded by the temporal pattern of the response of auditory neurons, and frequency by both response pattern and by which neurons respond among the many thousands in the auditory nervous system. A consequence of the interrelationship between time and frequency, and the direct coding of each aspect by the auditory system, is that studies of auditory temporal factors must ensure that the basis of measured effects is indeed temporal. Thus, studies of temporal factors often use stimuli with wideband frequency content or include maskers to restrict the influence of unintended frequency variations.

A further effect that ties together time and frequency is their inverse relationship in resolution. That is, enhancement of frequency selectivity results in a sluggish temporal response, and a rapid temporal response results in broad or poorly tuned frequency selectivity. The complexity of the auditory system no doubt arose partly to meet the conflicting demands of temporal and frequency resolution.

Variations in the time waveform of sound can generally be described in terms of modulation. *Modulation* simply refers to some pattern of change conveyed by a carrier. For auditory processing, the relevant variations are amplitude modulation (AM) and frequency modulation (FM). At slow rates of variation, listeners can follow the modulation, with AM perceived as fluctuation in loudness and FM as fluctuation in pitch. At faster rates, both types of modulation change the stimulus quality, timbre, or pitch, often adding a richness or roughness to the sound. Modulated stimuli play an important role in assessing temporal resolution and discrimination abilities, and also temporal-pattern processing.

Temporal Resolution

In a general sense, temporal resolution refers to how quickly a system is able to respond to stimulus changes. For auditory stimuli, these are most often changes in loudness, pitch, timbre, or source location. If the changes are not well resolved, the response of the auditory system represents some sort of averaging across the stimulus changes. Consequently, temporal resolution is often viewed as a measure of the minimum integration time of the auditory system.

A variety of procedures have been used to assess auditory temporal resolution. One simple procedure, termed *gap detection,* involves determination of the shortest silent interval that can just be detected when inserted in a sound burst or continuous sound. Gap detection thresholds can be as low as a few one thousandths of a second (i.e., several milliseconds [ms]). Roughly similar acuity estimates are obtained from other procedures that measure sensitivity to rapid changes in sound other than the introduction of a brief silent gap. A notable exception, however, is seen in measures of binaural temporal acuity. The just-detectable timing difference between the sounds arriving at the two ears can be as low as a few millionths of a second.

Related to gap detection are effects observed in temporal masking. Temporal masking refers to situations in which a target signal is either preceded or followed by a masking sound. Interference from the interaction between sequential sounds is a common aspect of many everyday listening situations. With temporal separation between the signal and masker in temporal masking paradigms, interference represents a failure of temporal resolution; that is, the system could not fully resolve the signal from the masker. In the case of forward masking in which the masker precedes the signal, two separate factors have been suggested as the source of the interference. The first is a persistence of the sensory effect evoked by the masker so that despite the acoustic temporal separation, the responses to the signal and masker are overlapping at some level of the auditory system. The second factor is neural adaptation in which the response to the masker adapts the system so that it is less responsive to presentation of the subsequent signal. Across various possible stimulus configurations, both factors can show some involvement in temporal masking.

The envelope of an acoustic signal describes how amplitude or level changes over time. As such, AM is modulation of the signal envelope. Gap detection measures sensitivity for a single envelope modulation, that is, the detection of the single brief silent interval or gap. In AM detection paradigms, the modulation occurs over the entire stimulus duration with the procedure determining the size of the just-detectable variation or fluctuation in the stimulus envelope. AM changes the perceived quality of a sound. At low AM rates that lead to slow level changes, listeners perceive the changes as fluctuations in loudness. At higher modulation rates at which the loudness variations are not resolved, AM results in a change in sound timbre often described as “motor-boating” or roughness.

AM detection procedures measure temporal resolution by showing the change in the ability to respond to envelope fluctuation as a function of the rate at which the fluctuation occurs. In psychoacoustics, this function is termed the *temporal modulation transfer function* (TMTF). With appropriate assumptions, a measure of temporal resolution can be derived from the TMTF. Consistent with other measures of temporal acuity, TMTFs estimate minimum-integration times of several milliseconds.

Temporal Integration

Temporal integration refers to the ability to summate response or information to affect perception. Results from detection experiments indicate that for durations of up to several hundred milliseconds, lower signal levels are needed for detection as signal duration is increased. These integration times in the tens or hundreds of milliseconds are several orders of magnitude greater than are those derived from measures of temporal resolution. Along with affecting signal detection, other effects of stimulus duration are commonly observed both in task performance, which can improve with stimulus duration (e.g., frequency discrimination ability), and increase in the strength of percept (e.g., loudness). In some cases, the change in task performance or percept with stimulus duration is not a result of actual temporal integration. Changing duration affects the frequency content of a sound with variation in frequency content the basis of some duration effects.

Several approaches have been used to reconcile the time constants of resolution and integration experiments. Based on the complex structure of the auditory system, one approach assumes that differing estimates reflect task-dependent processing at different levels of the system. Alternative models attempt to account for both temporal resolution and integration with a single processing structure. For example, the "multiple look" model proposes multiple sampling of the short time-constant process of temporal acuity. By combining information across multiple looks, the model is able to account for results from integration experiments.

Another type of information integration is observed with modulated stimuli. In these cases, the integration is across frequency regions. An example of a typical stimulus configuration would combine a low- and high-pitched tone, both having the same pattern of AM. This common modulation across frequency regions can aid detection of a signal added to the mixture. The improvement in detection resulting from common or coherent modulation is called comodulation masking release (CMR). If the task requires AM detection, modulation added to a different frequency region can hinder performance, with this effect termed modulation detection interference (MDI). For both CMR and MDI, information concerning temporal modulation is combined across frequency regions to influence performance.

Auditory Segregation and Pattern Processing

In some measures of temporal acuity, the onset of specific frequency components is delayed with the task to detect the presence of the asynchrony. Delaying the onset of a single component of a complex is referred to as asynchronous gating. With asynchronous gating at roughly 10 times the detection threshold (i.e., roughly 20 ms), listeners detect the delay and can report the temporal order at which frequency components are turned on. This judgment requires ability to perceptually segregate, or "hear out," individual frequency components from a complex sound. The effect of asynchronous gating can be quite robust, among the strongest of any of the stimulus factors that influence perceptual segregation. Auditory processing of asynchrony and temporal order are crucial in many everyday listening situations. Asynchrony is a common marker that distinguishes one sound source from another, and prescribed temporal order is one of the major acoustic characteristics of speech.

A gating asynchrony among stimulus components can also lead to the perception of sequential auditory events. The perception of multiple auditory events as a coherent sequence, in essence the output of a single sound source, is referred to as streaming. Common examples of auditory streaming are the ability to follow a single melody from a mixture of musical sounds, or attributing consecutive speech sounds to a single talker. A general characteristic of streaming is that pattern recognition is superior for elements that have arisen from a common source in contrast to multiple real or implied sources. This characteristic is often observed in experiments that evaluate whether a repeating series of sounds forms a single or multiple stream sequences. When using tonal stimuli, the perception can be one of hearing one or several concurrent melodies. In the simplest case, the stimulus is a repeating series of two pure tones with the frequency separation of the tones and the repetition rate being the primary experimental variables. Increasing the value of either of these variables increases the segregation of the two tones into separate streams. Long before experimental validation, awareness of these effects was used in contrapuntal music to effect segregation of melodic lines. In experimental work, the absolute values of the frequency separation and repetition rate at the threshold for stream segregation can vary with the metric used to determine threshold and the requirements of the task. Dependent on stimulus parameters, listeners may report an ambiguous percept that switches between stream segregation and sequence integration. In many studies, results also indicate a buildup of the percept of stream segregation over time. Collectively, results obtained from both temporal-order and streaming experiments highlight that time is the dimension that defines auditory patterns, making temporal processing a fundamental property of auditory perception.

Stanley Sheft

See also Audition: Pitch Perception; Auditory Frequency Selectivity; Auditory Masking; Auditory Scene Analysis; Auditory Thresholds; Sound Stimulus

Further Readings

Bregman, A. S. (1990). *Auditory scene analysis.* Cambridge: MIT Press.

de Boer, E. (1985). Auditory time constants: A paradox? In A. Michelsen (Ed.), *Time resolution in auditory systems* (pp. 141–158). Berlin: Springer.

Plomp, R. (1964). Rate of decay of auditory sensation. *Journal of the Acoustical Society of America, 36,* 2 77–282.

Viemeister, N. F., & Plack, C. J. (1993). Time analysis. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Human psychophysics* (pp. 116–154). New York: Springer.

AUDITORY FREQUENCY ANALYSIS, NEURAL

Speech, music, and every other sound one hears can be described by the pattern of amplitude at different frequencies. This pattern, which is called the sound's spectrum, distinguishes one sound from others. In human speech, for example, the vowel in the word *heat* sounds different from the vowel in *hat* because amplitude peaks occur at different frequencies in the spectrum of the first vowel. Frequency analysis, described in this entry, refers to the ability to process different regions of the spectrum separately. This ability allows a person to discriminate two tones of different frequencies, but more importantly, it allows the locations of the amplitude peaks in the spectrum of any sound to be encoded and represented in the brain. Frequency analysis begins in the cochlea, where the spectrum of the acoustic signal is converted to a representation called a *place code* for frequency. The place code for frequency that is created in the cochlea is preserved as the signal is represented in the responses of neurons and processed in the brain. Throughout the auditory system, the representation of frequency by a neural place code is the predominant organizational principle.

Sound arriving at the eardrum is transmitted through the middle ear into the cochlea, where it causes a long, narrow structure called the basilar membrane to vibrate. Because of the way the stiffness of the basilar membrane varies along its length, it responds to a sine wave stimulus with a traveling wave whose maximum displacement occurs at a location that varies systematically with the frequency of the tone. Each location along the basilar membrane responds best to a restricted range of frequencies; that is, each location is "frequency-selective." The frequency selectivity of any single location can be described by its tuning curve, which shows the pattern of thresholds as a function of stimulus frequency. Each threshold is defined as the level of a sine-wave stimulus that displaces the basilar membrane by a fixed amount. Tuning curves are V-shaped, and the frequency for which the threshold is lowest is called the *characteristic frequency* (CF). Thresholds increase rapidly for frequencies above CF, and increase more slowly for frequencies below CF. This pattern of thresholds means that each location responds to some frequencies but not others.

The mechanical frequency analysis accomplished by the basilar membrane is enhanced by the action of outer hair cells (OHCs). When stimulated by low-amplitude vibration of the basilar membrane, OHCs generate a mechanical response of their own. The effect of this "electromotile" response is called the cochlear amplifier. The feedback provided by the cochlear amplifier increases sensitivity to low-amplitude sounds near the CF and improves frequency selectivity.

The next step in frequency analysis occurs when the vibration of the basilar membrane is converted into neural activity. Motion of the basilar membrane stimulates inner hair cells (IHCs). IHCs in turn stimulate auditory nerve fibers (ANFs), the neurons that carry information about sound from the cochlea to the brain. Because each ANF receives input from a single IHC, it is also frequency-selective, with a tuning curve that is essentially the same as the tuning curve of the basilar membrane location that it innervates. A sine-wave stimulus will activate a subset of ANFs that has CFs that are close to the frequency of the tone. When the frequency of the tone changes, it activates a different subset of ANFs. This change in the place of activation provides the neural basis for frequency discrimination. Of course, most sounds are more complex than sine-wave tones. For a sound like one of the vowels mentioned previously, ANFs whose CFs are close to the frequencies of peaks in the vowel's spectrum will be activated most strongly. When the sound's spectrum changes, so that spectral peaks occur at different frequencies, the locations of peaks of neural activation will also change. This description leaves out many details, especially the contributions of cochlear nonlinearities, but in general, the pattern of activation as a function of place provides the basis for neural frequency analysis.

ANFs also convey information about frequency in their temporal discharge patterns. For frequencies below about 4 kilohertz (kHz), ANFs respond with a temporal pattern that is related to the waveform of the stimulus. For example, consider an ANF responding to a sine-wave tone whose frequency is 333 hertz (Hz). The waveform repeats every 0.003 seconds (3 msec). The fiber will respond at precisely timed intervals determined by that period: 3 msec, and integer multiples of 3 msec (6 msec, 9 msec, 12 msec, etc.). Now suppose that the frequency of the stimulus changes to 250 Hz. The period changes to 4 msec, and the interval between responses also changes, to 4 msec and integer multiples of 4 msec. Sounds that are more complex than sine waves are also represented with a precise temporal pattern. This response property, called phase-locking or discharge synchrony, contributes to frequency analysis, especially in the presence of noise.

The place code for frequency created in the cochlea is preserved in the central auditory system. The tuning curves and CFs of central auditory neurons are determined by their direct and indirect connections to auditory nerve fibers. In every structure in the main ascending auditory pathway, neurons are arranged so that CF changes systematically as a function of location. In the brain, this is referred to as tonotopic organization. Tonotopic organization allows the neural frequency analysis that starts in the cochlea to be completed.

Donal G. Sinex

See also Audition: Pitch Perception; Auditory Frequency Selectivity; Auditory Processing: Peripheral; Auditory Receptors and Transduction; Tuning Curves

Further Readings

Dallos, P. (1992). The active cochlea. *Journal of Neuroscience, 12,* 4575–4585.

Liberman, M. C. (1978). Auditory nerve responses from cats raised in a low-noise chamber. *Journal of the Acoustical Society of America, 63,* 442–455.

Ruggero, M. A., Narayan, S. S., Temchin, A. N., & Recio, A. (2000). Mechanical bases of frequency tuning and neural excitation at the base of the cochlea: Comparison of basilar-membrane vibrations and auditory-nerve-fiber responses in chinchilla. *Proceedings of the National Academy of Science USA, 97,* 11744–11750.

AUDITORY FREQUENCY SELECTIVITY

Our world is filled with acoustic energy. Sound waves radiate from different sources, bouncing off walls, floors, and other objects, and intermingling before finally reaching our ears. What the ear receives is a complex mixture of sources and reflections, all combining to form a single pressure wave that causes the eardrum to vibrate. How do we segregate such a mixture, and how do we identify and recognize the sources? This question, which forms part of the well-known cocktail party problem, is still not completely solved. However, one contributing factor is our ability to hear and segregate different frequencies within the sound mixture. This *frequency selectivity,* described in this entry, is established at an early stage of auditory processing and is maintained throughout the auditory pathways, right up to the primary auditory cortex.

Frequency Selectivity in the Cochlea

When sound vibrations reach the cochlea via the eardrum and the middle ear bones (ossicles), the vibrations disperse as "traveling waves" along one of the partitions within the cochlea, known as the basilar membrane. How far along the basilar membrane a wave travels depends mainly on its frequency, with high-frequency waves reaching their peak early and low-frequency waves peaking further down along the basilar membrane. In this way, a complex waveform is broken down into its constituent simple frequencies so that the energy at each frequency produces its own peak of vibration along the basilar membrane. The pattern of vibration along the basilar membrane in response to a sound is called an *excitation pattern.*

The frequency selectivity of the cochlea is limited: As two tones become closer in frequency, their peaks of vibration along the basilar membrane will also become closer, and at some point the two peaks will merge to form an excitation pattern with just a single peak. The limits of cochlear frequency selectivity are determined by the mechanical properties of the basilar membrane, as well as by the functioning of the outer hair cells, which sit on top of the basilar membrane, within the organ of Corti, and amplify quiet sounds and sharpen the tuning of the cochlea. Damage to the outer hairs results in a reduction in cochlear tuning, and a loss of frequency selectivity. In other words, when the outer hair cells are not functioning properly, the excitation patterns in the cochlea become broader, making it more difficult to separate the responses to different frequencies.

Perceptual Measures and Consequences of Frequency Selectivity

Masking has been a popular way to probe perceptual frequency selectivity in the laboratory, but it is also an everyday occurrence—think of holding a conversation in a noisy bar, or of not hearing someone because you have the vacuum cleaner on. Masking can occur when the excitation pattern produced by the masker overlaps with, and swamps, the excitation produced by the target. If auditory frequency selectivity were infinitely sharp, then no masking would occur, as long as the frequencies of the masker did not overlap completely with those of the target. Thus, the extent to which masking does occur can tell us about the limits of frequency selectivity. Some rules of thumb have emerged from research into frequency selectivity. In general, two frequencies have to be separated by at least 10 to 20% to be perceptually segregated, at least in the range of frequencies between about 500 and 6,000 Hz, which are arguably the most important for speech perception. The ability of one sound to mask another diminishes as the frequency separation between them increases and, at high sound levels, low-frequency sounds mask high-frequency sounds more readily than the reverse, in what is known as the upward spread of masking.

Frequency selectivity can be determined physiologically in the auditory nerve, or along the basilar membrane, by measuring the range of frequencies that elicit a given response. The measurement of these physiological "tuning curves" requires for the most part invasive procedures and so cannot be performed in humans. However, the few studies that have investigated both perceptual and physiological frequency selectivity in the same species have found good correspondence, meaning that perceptual frequency selectivity is likely determined in the cochlea. Hearing loss often results from a functional loss of the outer hair cells, leading to poorer frequency selectivity in the cochlea. Perceptual studies of hearing-impaired people have found that their frequency selectivity is poorer than normal. This decrease in frequency selectivity may underlie many common complaints of people with hearing loss, such as an impaired ability to understand speech in noisy backgrounds.

The limits of frequency selectivity can also be put to positive use in certain practical applications. For instance, the popular audio coding scheme known as MP3 relies on humans' limited frequency selectivity to "hide" coding noise in frequency regions where it is masked by the audio signal. In MP3 coding, the original audio signal is split up into frequency "sub-bands" and the resolution of the signal within each sub-band is degraded, which is the equivalent of adding noise. Because the added noise is limited to the same frequency region as the signal, engineers can be sure that the signal will perceptually mask the added noise, so long as its level remains well below that of the signal. By making clever use of frequency selectivity and masking, the designers of the MP3 and related systems have been able to reduce the data rate by a factor of nearly 10 without any noticeable loss in audio quality.

Andrew Oxenham

See also Audition; Auditory Frequency Analysis, Neural; Auditory Masking; Auditory System: Structure; Auditory Thresholds

Further Readings

Moore, B. C. J. (1995). Frequency analysis and masking. In B. C. J. Moore (Ed.), *Handbook of perception and cognition: Vol. 6. Hearing* (pp. 161–205). Orlando, FL: Academic Press.

Shera, C. A., Guinan, J. J., & Oxenham, A. J. (2002). Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proceedings of the National Academy of Sciences USA, 99,* 3318–3323.

AUDITORY ILLUSIONS

The sounds we perceive do not always correspond to those that are presented. When such a mismatch occurs, we are experiencing an *auditory illusion.* These illusions show that the auditory system does not faithfully transmit the sound information as it arrives at our ears, but alters and reorganizes this information in various ways, as described in this entry.

The Precedence Effect

Our hearing mechanism has evolved an ingenious mechanism for minimizing problems caused by echoes in the environment. Instead of correctly perceiving a set of overlapping sounds, each coming from a different location in space, we obtain the illusion of a single sound that appears to be coming from its original source. This phenomenon was first discovered by the 19th-century physicist Joseph Henry. To demonstrate this effect, the listener is seated in front of two loudspeakers, with one to his left and the other to his right. A single stream of speech is presented through both loudspeakers; however, the signal at the right loudspeaker is delayed relative to the left one. When the sounds coming from the speakers differ in onset by less than about 30 milliseconds (ms), the listener perceives the sound as coming only from the left loudspeaker. The right loudspeaker contributes to the loudness and liveliness of the sound, but appears to be completely silent. When the offset exceeds the critical time limit, two distinct streams of sound are correctly heard as coming from separate loudspeakers.

The Octave Illusion

The octave illusion was discovered by Diana Deutsch in 1973 and is experienced with tones that are presented via stereo headphones. Two tones an octave apart are alternated repeatedly at a rate of four tones per second. The identical sequence is presented to both ears simultaneously; however, the tones are offset in time such that when the right ear receives the high tone the left ear receives the low tone; and vice versa (see Figure 1, left panel).

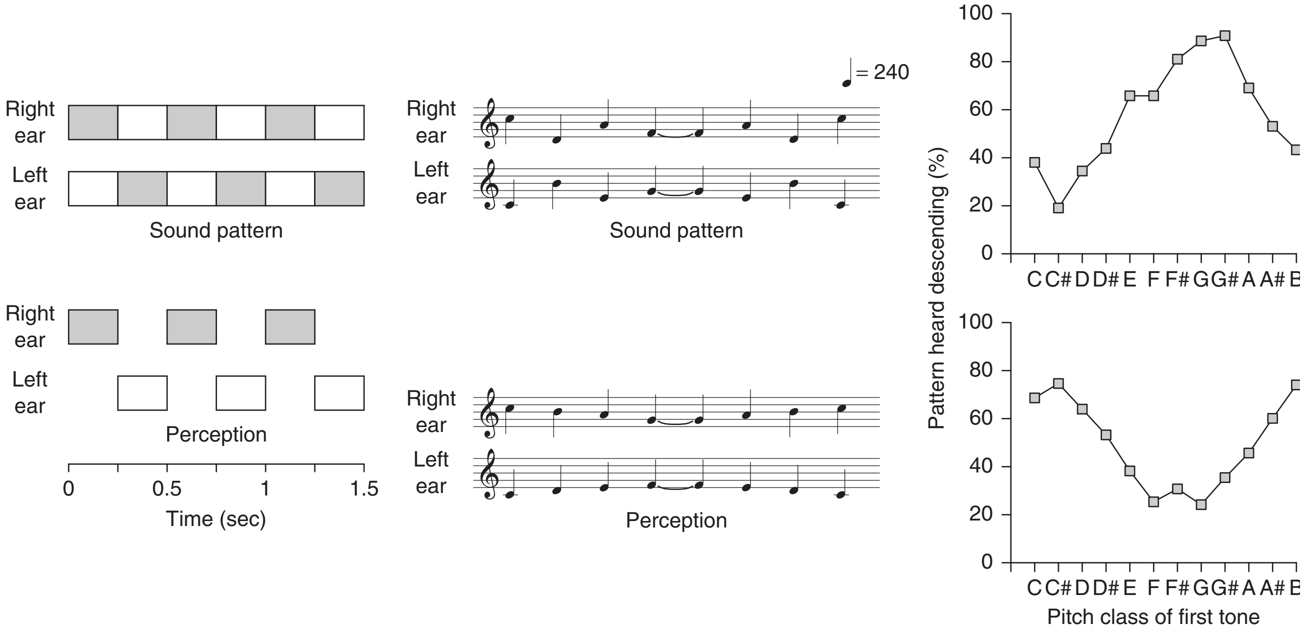
Despite its simplicity, this pattern is almost never heard correctly, and instead produces a number of illusions. Many people hear a single tone that alternates from ear to ear, while its pitch simultaneously switches back and forth between high and low. So it appears that one ear is receiving the sequence "high tone—silence—high tone—silence" while the other ear is receiving the sequence "silence—low tone—silence—low tone." Even more curiously, when the earphone positions are reversed most people hear exactly the same thing: The tone that had appeared in the right ear still appears in the right ear, and the tone that had appeared in the left ear still appears in the left ear. As a further surprise, right-handers tend to hear the high tone as on the right and the low tone as on the left, regardless of how the earphones are positioned; yet left-handers vary considerably in where the high and low tones appear to be coming from. The illusion is hypothesized to result from the use of incompatible cues to determine *what* sounds are being presented, and *where* each sound is coming from. The strong association with handedness indicates that the ways in which the octave illusion is perceived reflect differing patterns of brain organization.

The Scale Illusion

The scale illusion was discovered by Deutsch in 1973 and is best experienced through stereo headphones. The pattern that produces this illusion consists of a musical scale, played simultaneously in ascending and descending form. The successive tones of the scale alternate from ear to ear such that when a tone from the ascending scale is in the right ear a tone from the descending scale is in the left ear; and vice versa. This sequence of tones is played continuously without pause (see Figure 1, center panel).

As with the octave illusion, the scale pattern gives rise to a number of illusions that vary with the handedness of the listener. Most commonly, a melody corresponding to the higher tones is heard as coming from one earphone, with a melody corresponding to the lower tones from the other one. Right-handers tend to hear the higher tones as on their right and the lower tones as on their left. And as with the octave illusion, when the earphone positions are reversed the right ear continues to hear the higher tones, and the left ear the lower ones. However, left-handers as a group are less likely to localize the tones in this way. As with the octave illusion, the handedness correlates with the way the scale illusion is perceived, indicating that this illusion reflects differing patterns of brain organization.

Variants on the scale illusion can be produced with other melodic patterns, and by natural instruments in concert halls, so that on listening to complex orchestral music, individual members of an audience may well be hearing certain musical passages in different ways.



(a) Octave Illusion **(b) Scale Illusion** **(c) Tritone Paradox**

Figure 1 Auditory Illusions

Sources: (a) Adapted from Deutsch (1974). (b) Adapted from Deutsch (1975). (c) Adapted from Deutsch (1995).

Notes: (a) The pattern that produces the octave illusion, and the percept most frequently obtained. Filled boxes indicate tones of 800 hertz (Hz) and unfilled boxes indicate tones of 400 Hz. (b) The pattern that produces the scale illusion, and the percept most frequently obtained. (c) The tritone paradox as perceived by two different listeners.

The Glissando Illusion

The glissando illusion was created by Deutsch in 1995. It is best heard when the listener is seated in front of two stereophonically separated loudspeakers, with one to his left and the other to his right. The pattern producing the illusion has two components: a synthesized oboe tone of constant pitch, and a sine wave whose pitch glides up and down. These two components are presented simultaneously, and switch repeatedly between the loudspeakers at four cycles per second such that when the oboe tone is coming from the speaker on the left, the glissando is coming from the speaker on the right, and vice versa.

On listening to this illusion, the oboe tone is heard correctly as switching between the loudspeakers. However, the segments of the glissando appear to be joined together quite seamlessly, so that a single, continuous tone is heard that appears to be moving around in space in accordance with its pitch motion. Handedness correlates again emerge where the higher and lower portions of the glissando appear to be located: Right-handers tend strongly to perceive the glissando as moving slowly in space between a position to the left when its pitch is low, and to the right when its pitch is high. Yet left-handers vary considerably in where the higher and lower portions of the glissando appear to be located.

The scale and glissando illusions can be partly explained by assuming that the listener adopts the most plausible interpretation relative to the environment. For the scale illusion, it is extremely unlikely that two jerky but overlapping pitch patterns should be coming from different positions in space. We therefore reorganize the tones perceptually, so that the higher tones appear to be coming from one location and the lower tones from another. For the glissando illusion, it is extremely unlikely that a sound that changes smoothly in frequency should be switching abruptly between two different spatial locations. It is far more probable that such a sound is coming from a single source that is either stationary or that is moving slowly through space. So we perceptually reorganize the gliding tone accordingly.

Pitch Circularities

Tones whose frequencies are in a ratio of 2:1 are said to be in octave relation, and they are in a sense perceptually equivalent. This is acknowledged in the system of notation of the Western musical scale. The core of this scale consists of 12 tones, which correspond to the division of the octave into equal steps, called semitones, and each tone is given a name (C, C#, D, D#, E, F, F#, G, G#, A, A#, and B). The entire scale, as it ascends in height, consists of the repeating presentation of this sequence of note names across successive octaves. Pitch can therefore be described as varying along two separable dimensions: The dimension of *pitch height* corresponds to the position of the tone from high to low, and the circular dimension of *pitch class* (corresponding to note name) corresponds to the position of the tone within the octave.

In the early 1960s, Roger Shepard showed that the two dimensions of pitch can be separated by employing computer-produced tones whose pitch classes are clearly defined, but whose heights are ambiguous. This can be achieved, for example, with a bank of 12 tones whose names correspond to those of the musical scale, but each of which consists only of components that are related by octaves. For example, one such tone would consist only of components C_2, C_3, C_4, . . . ; such a tone would clearly be heard as a C, but its octave placement would be ambiguous.

Shepard found that when two such octave-ambiguous tones were played in succession, people heard either an ascending pattern or a descending one, depending on which was the shorter distance between the tones along the pitch class circle. This enabled Shepard to create an illusion of pitch circularity: When a bank of such tones is played that traverses the pitch class circle clockwise in semitone steps (C, C#, D, and so on), it is heard as eternally ascending in pitch. When it traverses the circle counterclockwise (C, B, A#, and so on), it is heard as eternally descending instead. Soon after this, Jean-Claude Risset generated analogous pitch circularities using a similar algorithm but with gliding tones.

Pitch circularities are not confined to tones consisting only of components that are related by octaves. In 2008, Diana Deutsch, Kevin Dooley, and Trevor Henthorn produced circular banks of tones each of which constituted a full harmonic series. This was achieved by manipulating the relative amplitudes of the odd and even harmonics of the tones.

The Tritone Paradox

Demonstrations of pitch circularity raise the question of what happens when pairs of octave-ambiguous tones are presented that stand in opposite positions along the pitch class circle, so that the same distance between them can be traversed in either direction. Such tones form an interval called a tritone. Deutsch discovered in 1983 that when such tone pairs are presented, paradoxical differences in perception emerge (see Figure 1, right panel). For example, when presented with C followed by F#, one listener will clearly hear an ascending pattern; however, another listener may clearly hear a descending pattern instead. And when the pattern G# followed by D is presented, the first listener now hears a descending pattern and the second listener an ascending one. Furthermore, for any one listener, the pitch class circle is oriented with respect to height in a systematic way, so that tones in one region of the circle are heard as higher and tones in the opposite region as lower.

Such differences in perception correlate with the pitch range of the listener's speaking voice, and with the language or dialect to which listeners have been exposed. Strong differences were found, on a statistical basis, between listeners who had grown up in California and those who had grown up in the south of England. Such differences in perception of this musical pattern again indicate that individual members of an audience may hear certain musical passages in different ways.

Illusions of Auditory Continuity

Another class of illusions involves perceptually replacing missing portions of sounds that would otherwise be drowned out by noise. When we are conversing with someone in the street, for example, the noise of traffic masks out many portions of the speech we are listening to, and our perceptual system generates these missing portions, so that the speech sounds continuous. Richard Warren found that when listeners heard a recorded sentence in which a speech sound was deleted and replaced by noise, they were unable to tell which sound was missing, and could not locate the position of the noise in the sentence. The sound that listeners "reconstructed" changed depending on the context. When presented with the sentence "It was found that the *eel was on the ___," listeners tended to "hear" the missing sound depending on the last word. So, for example, when the last word was "table," listeners tended to hear the word "meal," and when the last word was "shoe," they tended to hear the word "heel." Similar effects have been found on listening to music. When a note in a repeating scale is deleted and replaced by a noise burst, listeners reconstruct it perceptually. Also, as discovered by Albert Bregman, when a gliding tone is sounded, and a portion of the glide is removed and replaced by a loud noise burst, the listeners reconstruct the missing portion of the glide.

Diana Deutsch

See also Audition: Cognitive Influences; Audition: Pitch Perception; Auditory Imagery; Auditory Localization: Psychophysics; Auditory Scene Analysis; Feature Integration Theory

Further Readings

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge: MIT Press.

Deutsch, D. (1974). An auditory illusion. *Nature, 251,* 307–309.

Deutsch, D. (1975). Two-channel listening to musical scales. *Journal of the Acoustical Society of America, 57,* 1156–1160.

Deutsch, D. (1986). A musical paradox. *Music Perception, 3,* 275–280.

Deutsch, D. (1995). *Musical illusions and paradoxes.* La Jolla, CA: Philomel Records.

Deutsch, D. (Ed.). (1999). *The psychology of music* (2nd ed.). San Diego, CA: Academic Press.

Deutsch, D., Dooley, K., & Henthorn, T. (2008). Pitch circularity from tones comprising full harmonic series. *Journal of the Acoustical Society of America, 124,* 589–597.

Risset, J. C. (1969). Pitch control and pitch paradoxes demonstrated with computer-synthesized sounds. *Journal of the Acoustical Society of America, 46*(1A), 88.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America, 36,* 2345–2353.

Wallach, H., Newman, E. B., & Rosenzweig, M. R. (1949). The precedence effect in sound localization. *American Journal of Psychology, 62,* 315–336.

Warren, R. M. (1983). Auditory illusions and their relation to mechanisms normally enhancing accuracy of perception. *Journal of the Audio Engineering Society, 31,* 623–629.

AUDITORY IMAGERY

Auditory imagery refers to the common experience in which people report that they "hear" a voice, melody, or other sound, in their "mind's ear," all in the absence of an actual acoustic stimulus. In some cases, this experience is deliberate (and so someone can, if he or she chooses, try to imagine what his or her mother's voice sounds like, or what a particular musical performance sounds like). In other cases, the image arises spontaneously, and, indeed, people sometimes complain that they cannot avoid hearing a melody (or a snippet of a melody) over and over—a maddening experience sometimes given the striking label of "having an ear-worm."

Auditory images are not hallucinations—people experiencing the images can tell that the images are "in their head," and not a real sound. Nonetheless, the experience of "hearing" the imagined sound does resemble the experience of hearing an actual sound. Subjectively, the imagined sound seems to have a clear pitch, duration, and timbre, just as an actual sound would. Functionally, the image seems to provide direct information about these attributes, suggesting that the image truly does depict the sound, rather than merely describing or referring to the sound. This entry describes imagined pitch, duration, and timbre; differences between sounds and auditory images; enacted auditory images; the neural substrate of auditory imagery; and memory for enormously familiar sounds.

Imagined Pitch and Imagined Duration

Research on auditory imagery has taken several different paths. One line of research has sought to confirm the subjective sense that auditory images do directly represent a sound's pitch. One series of experiments, for example, asked participants to imagine a specific pitch, and then to detect a faint tone either of the same pitch or different. Participants' performance was better when the tone they were trying to detect was the same pitch as the one they were imagining—suggesting both that the image had accurately represented the pitch and that this imagined pitch primed the processes of actual hearing. A different series of experiments asked participants to imagine a particular melody and then to hum the starting pitch of the melody they were thinking about. Two days later, participants returned to the lab and did the same task. Across this two-day interval, the data show remarkable consistency in the pitch that each participant hummed for each song—indicating both that participants imagined the song at a specific pitch and that they were consistent in the pitch they chose for each melody.

A different line of research confirms the subjective sense that images are stretched out in time in just the way actual sounds are. In one study, participants were given specific words from a song's lyrics (for example, *can* and *by* from "The Star-Spangled Banner"), and had to judge whether the pitch of the note accompanying the second word (*by,* the seventh beat of the phrase) was higher or lower than the pitch of the note accompanying the first word (in this case, the third beat). The data showed that the time needed to make this judgment increased in a regular fashion if the first note was further from the song's start, and also if the second note was further from the first. It would seem, then, that participants perform this task by "playing" the melody "in their heads," starting with the melody's actual start—and the more notes they have to "play" to make their judgment, the more time needed.

Imagined Timbre

Still other research confirms that auditory images accurately depict a sound's timbre (i.e., the quality of sound, with other properties, such as pitch or loudness, held constant, that distinguishes one sound source from another; for example, the quality that distinguishes a flute and a clarinet each playing the same pitch). For example, in one study, some participants had to rate the similarities of a series of actually presented sounds that differed in timbre. (In one version of the procedure, the sounds were those of orchestral instruments; in another version, the sounds were environmental sounds.) Other participants had to *imagine* these sounds, and also rated the similarities between the sounds. The statistical summary of these data relied on a procedure termed *multidimensional scaling*. In this procedure, the individual sounds are represented by distinct points in space, and the similarity ratings are interpreted as "distances" among these points; the analysis asks how the sounds must be positioned to create the complex pattern of "distances," from one sound to the next, revealed in participants' pair-wise ratings.

The multidimensional space produced by this analysis, based on the similarity ratings for the imagined sounds, closely resembled the space based on the ratings for the actual sounds, indicating that all the pair-wise relationships between sounds are well preserved in imagery, which implies in turn that the timbres were represented in a fashion that is at least consistent and, more strongly, true to the actual sounds.

Differences Between Sounds and Auditory Images

Other results, however, suggest some complications in the comparison between imagined sounds and actual sounds. For example, actual sounds necessarily have some specific loudness; findings suggest, however, that auditory images may be indeterminate in their loudness. In addition, auditory images may not be purely "auditory," because some data suggest that imagination may often represent complex, multimodal events, containing both auditory and visual aspects. As a result, some judgments about auditory images may be guided by information about other aspects of the represented events. Thus, for example, participants imagining familiar *voices* may create an image that also depicts the familiar *face* of the person speaking, and judgments about the voice—such as whether it is similar to some other voice—may be shaped by information about the visual qualities of the face.

Auditory images are also distinct from sounds in another way: Sounds exist in a fashion that is independent of any perceiver's interpretation—how the perceiver parses the sound-stream (i.e., determines where one sound stops and the next begins, and decides which sounds are parts of a larger acoustic event), or interprets its meaning. Auditory images, in contrast, seem to exist only in the context of the perceiver's interpretation; according to some authors, the image is accompanied by a *perceptual reference frame* that specifies how the sound is organized and understood. As a result of this reference frame, the meaning of the imaged sound, for the person holding the image, may be quite specific and relatively rigid. In one study, for example, some participants heard the word *stress* being uttered again and again and again; within a few seconds, their perception shifted, so they heard the sequence first as repetitions of *stress,* then of *rest,* then of *tress*—reflecting apparently spontaneous shifts in how the sound stream was parsed. If these repetitions of *stress* were imagined, however, the parsing of the stream was apparently specified within the perceptual reference frame that accompanied the image, and so these shifts in interpretation did not occur, and the (imagined) sequence was understood as a rigidly unchanging repetition of the initial word, *stress.*

Enacted Auditory Images

In some settings, however, participants seem able to escape the limits set by the perceptual reference frame. Specifically, in some tasks, participants seem spontaneously to supplement an imagined stimulus with silent subvocalization (and so, in addition to imagining the sounds of "stress, stress, stress," they might silently mime the speaking of this sequence). This combination of imagery plus subvocalization produces what some authors call enacted imagery, and enacted images (unlike "pure" images) seem readily re-interpreted (and so enacted images of "stress-stress-stress" are easily re-parsed).

The role of enactment is evident in a variety of other tasks that rely on auditory imagery. In one study, participants were given letter and number strings, and had to decide what these strings would sound like if pronounced aloud. (X-T-C, for example, would be *ecstasy;* N-L-S-S would be *analysis.*) Participants could easily do this task if given no other constraints. If, however, participants were blocked from subvocalizing (e.g., by asking them, silently and repeatedly, to mouth "ta-ta-ta" while doing the task), their performance dropped markedly. If participants were in a noisy environment (so that they couldn't "hear" what they were saying to themselves), performance was also quite low. It appears, then, that participants perform this task, in essence, by talking to themselves, and then "listening" to what they have (silently) pronounced. Put differently, this task, apparently relying on auditory imagery, seems instead to rely on a mix of auditory imagery and motor activity (the enactment), and if either the audition or the motor action is prevented, performance drops.

Neural Substrate of Auditory Imagery

Still another line of research has examined the brain mechanisms that support (and shape) auditory imagery. Both positron emission tomography and functional magnetic resonance imaging studies indicate that auditory imagery for familiar melodies depends on activation in the right auditory association cortex and activation in the frontal cortex. Tasks requiring judgments about musical timbres seem to activate both primary and secondary auditory areas, especially in the right hemisphere. These studies indicate considerable neural overlap between the brain areas needed for auditory imagery and those needed for the actual hearing of overt sounds (a pattern that parallels the findings for visual imagery and vision), underscoring the strong resemblance between imagery and perception. More specifically, we can see that imagery and perception resemble each other subjectively (images "feel like" hearing), functionally (e.g., in the information they make prominent), and biologically. In addition, activation of the supplementary motor area (SMA) also appears to be involved in the generation of auditory imagery, again suggesting a role for motor codes in auditory imagery.

Memory for Enormously Familiar Sounds

Finally, yet another line of research has examined imagery for materials that are enormously familiar—such as a favorite song that someone has heard countless times. Evidence suggests that when someone imagines this sort of familiar song, the person's imagination is remarkably faithful to the original: The song is imagined in the same key as the original, and probably in the right octave, and at essentially the right tempo. Apparently, then, the person's memory for this familiar auditory event is quite accurate, and the reproduction of the melody, in imagery, preserves the acoustic qualities of the original.

Even with these diverse lines of research, however, many questions about auditory imagery remain; indeed, far less is known about auditory imagery than about visual imagery. For example, subjective reports suggest that people differ in how rich or detailed their auditory images are, and in how often they experience spontaneous images. Little is known about these points, however, including the fundamental issue of whether these differences can be taken at face value, or whether, instead, people are quite uniform in their auditory imagery and differ only in how they describe their images. Or, as a different example, little is known about how (or when, or whether) people use their auditory images. There are many indications that visual images are often useful (as aids to memory, or to problem solving); the history of scientific discovery or of artistic innovation suggests many steps forward inspired by (or created through) visual images. Whether a similarly rich set of functions can be documented for auditory imagery remains among the many points in need of further research.

Daniel Reisberg

See also Absolute Pitch; Audition; Music Cognition and Perception; Visual Imagery

Further Readings

Halpern, A., Zatorre, R., Bouffard, M., & Johnson, J. (2003). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia, 42,* 1281–1292.

Kraemer, D., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery: Sound of silence activates auditory cortex. *Nature, 434,* 158–159.

Levitin, D. (1994). Absolute representation in auditory memory: Evidence from the production of learned melodies. *Perception & Psychophysics, 56,* 414–423.

Reisberg, D. (2000). The detachment gain: The advantage of thinking out loud. In B. Landau, J. Sabini, E.

Newport, & J. Jonides (Eds.), *Perception, cognition and language: Essays in honor of Henry and Lila Gleitman.* Cambridge: MIT Press.

Reisberg, D. (Ed.). (1992). *Auditory imagery.* Hillsdale, NJ: Lawrence Erlbaum.

AUDITORY LOCALIZATION: PHYSIOLOGY

The ability to locate the source of a sound is a fundamental property of normal hearing. Compared with the visual system, however, for which a relatively complete understanding of spatial processing has been achieved, at least with respect to the two-dimensional representation of visual space projected on the back of the retina, the study of neural coding of auditory space constitutes an altogether more complex problem. This is partly because the hearing organ of the inner ear—the cochlea—is arranged to represent the frequency of a sound, rather than the location of its source. This frequency tuning is known as *tonotopy,* with the highest frequency (pitch) sounds represented at the base of the coiled cochlea and the lowest at the apical end. Brain centers dedicated to processing hearing in the central nervous system are organized according to this tonotopic map, such that the preferred sound frequency to which neurons are found to respond changes systematically from low to high moving from one end of a sheet of brain tissue to the other. Neural sensitivity to auditory spatial cues only emerges several stages above the cochlea. This entry describes cues for location, neural sensitivity to binaural cues, maps of auditory space, the neural code for interaural time difference, and cortical codes for auditory space.

Cues for Localization

Sound localization is provided for by sensitivity to a combination of acoustic cues that arise from the interaction of sounds with the head. These include binaural (or "two-eared") cues of interaural time and level differences (ITDs and ILDs, respectively). ITDs arise because the ears are separated in space by the distance of the head. Sound sources directly in front (or behind) generate zero ITD because the sound sources are equidistant from the two ears; for sound-sources located to one side or the other, however, the sound arrives later at the further ear than at the nearer ear, generating an ITD. In humans, these ITDs are maximally in the order of several hundred microseconds (*millionths* of a second) for sources located directly to one side or the other. ITDs are useful for localizing low-frequency sounds (< 1,500 Hz; approximately the note G two octaves above middle C on the piano).

For frequencies higher than 1,500 hertz (Hz), the head creates an "acoustic shadow" such that the sound arriving at the ear further from the source is less intense than the sound arriving at the nearer ear, creating an ILD. As well as depending on the location of the source, the magnitude of the ILD depends on the frequency of the sound and the size of the head (ILDs are bigger at higher frequencies and bigger for bigger head sizes). This dichotomy of ITDs for low-frequency localization and ILDs for high-frequency localization is referred to as the *duplex theory* of sound localization.

A final cue for sound localization arises because of the convoluted shape of the outer ear, or pinna. The various nooks and crannies of the pinna lead to sound waves interacting with each other, with the result that sound energy is reduced at some frequencies and increased at others. This effect is different for different frequencies, depends on the location of the sound source, and leads to a different pattern of sound energy reaching the eardrum compared with that arriving at the outer ear. As well as contributing to the perception of source location above and below the horizon, these "spectral cues" are necessary for disambiguating sound sources located behind from those in front.

Neural Sensitivity to Binaural Cues

In mammals, the dichotomy suggested by the duplex theory is reflected in the brain structures that generate sensitivity to ITDs and ILDs. It is now well established that the medial superior olive (MSO) and lateral superior olive (LSO)—two aggregates of neurons in a region of the brainstem known as the superior olivary complex (SOC)—constitute the initial sites at which information from the two ears is brought together to create sensitivity to binaural hearing. Neurons in the MSO are sensitive to ITDs, and neurons in the LSO to ILDs.

Until recently, the most influential model for generating ITD sensitivity suggested that MSO neurons act as binaural coincidence detectors, neurons that respond maximally when nerve action potentials arrive simultaneously from the two ears. In this model, first developed by Lloyd Jeffress in 1948, each coincidence detector (MSO neuron) is innervated by axons from each ear, and these axons can vary in length. The purpose of these axonal "delay lines" is to compensate for ITDs generated by the spatial location of sound sources. This is illustrated in Figure 1 (see also color insert, Figure 1), where each neuron in the array is contacted by axons of varying length from each ear. So, for example, for a sound with zero ITD—generated by a source directly in front—signals from each ear enter the circuit simultaneously, and reach the middle coincidence detector (Number 3) simultaneously because the path length from each ear is identical. The firing of this detector, therefore, indicates that ITD was zero. Now consider what happens if the sound is located nearer the left ear, so that it is delayed in reaching the right ear. In that case, the signal from the left ear enters the circuit first, then the signal from the right ear, with a delay, such that the signals reach a neuron such as 4 or 5, which is closer to where the right-ear signal enters. The firing of this coincidence detector indicates that particular ITD and, by extension, the location of the source. Experimental recordings from a range of mammalian and avian species have demonstrated the existence of neurons that respond as coincidence detectors, being maximally activated by certain ITDs and minimally by others. This ability to compare the time of arrival of sounds at the two ears relies on the fact that the first neurons in the auditory pathway—the auditory nerve fibers (ANFs) that synapse with the sensory inner hair cells of the cochlea—generate action potentials only at specific times during each stimulus cycle, a process called *phase-locking*. Thus, by comparing the timing of action potentials arriving from each ear, it is possible to know the timing of the sound arriving at each ear and, hence, construct the location of the sound source. Phase locking of action potentials occurs for frequencies of sound up to a few kilohertz (kHz), and this upper limit is reduced between the ANFs and the MSO neurons. Hence, ITD processing is restricted to sounds with frequencies below about 1.5 kHz in mammals.

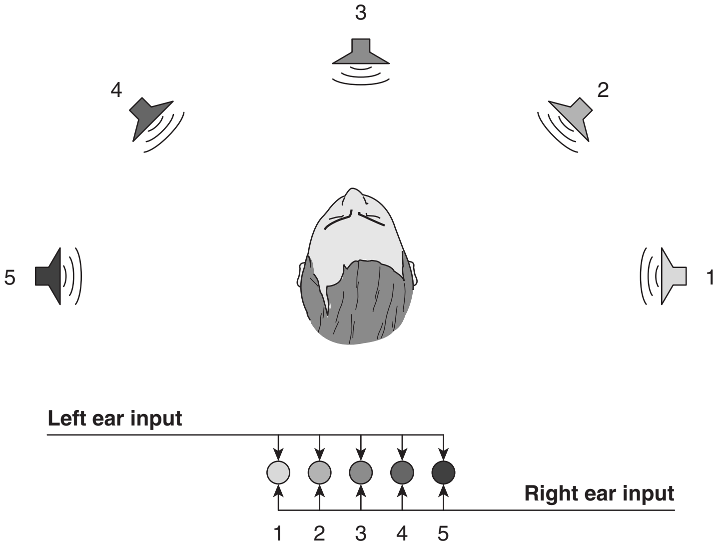


Figure 1 Axonal Delay Lines

Note: Each neuron in the array is contacted by axons of varying length from each ear.

Further, anatomical studies by Matsuka Konishi in barn owls, the "textbook" species for spatial hearing, suggests that ITD is mapped in an orderly manner across the neural tissue, with axonal delay lines favoring short ITDs at one end of the barn owl's equivalent brain-structure to the MSO, and delays lines favoring long ITDs at the other. This is mapped in the direction orthogonal (90°) to the tonotopic map, meaning that each sound frequency will contain a complete representation of ITDs. Evidence for such an arrangement of delay lines in the mammalian brain is equivocal, however, and mechanisms other than the Jeffress circuit for generating a neuron's preferred ITD have been suggested. One potential mechanism relates to the recent finding that MSO neurons are innervated by a significant, and temporally precise, source of inhibition. Studies in the laboratory of Benedikt Grothe in particular suggest that when the action of neural inhibition in the MSO is blocked by the drug strychnine, the difference in axonal conduction delay from each ear to the MSO is essentially zero. In other words, although neurons act as coincidence detectors, their preferred ITD is determined by neural inhibition, without which their preferred tuning would be for zero ITD (i.e., sounds straight ahead or behind).

The mechanism by which neural inhibition influences ITD tuning remains unclear. However, it is thought that because inhibitory inputs originate from the opposite side of the brain to the MSO and arrive directly on the soma (cell body) of a coincidence detector neuron, whereas excitatory inputs arrive from both ears distant from the soma—on the neuron's dendrites (processes)—the relative timing of the sound at each ear determines whether excitatory or inhibitory influences dominate the response. For one particular ITD, the inhibitory influence on the coincidence detection of the excitatory inputs is minimal, and the neuron responds maximally to this ITD. Regardless of the mechanism, the experimental demonstration that blocking inhibition shifts neurons' preferred ITD to zero indicates that the Jeffress model is only one means of generating a system of internal delays.

Generating neural sensitivity to ILDs is an altogether less complex affair than is the case for ITDs. It is generally accepted that ILD sensitivity arises first in the LSO, where neurons are innervated by excitatory synapses derived from the ear on the same side and the inhibitory synapse from the other. The generation of a significant head-shadow at higher sound frequencies for sound sources located to one side or the other results in either the excitatory or inhibitory influence dominating responses in each LSO. When the source is located closer to the excitatory ear, the LSO responds more vigorously, and less so as the source is moved relatively closer to the excitatory ear. Thus, the neural code for ILD, at least at the level of the LSO, is one in which the relative balance of activity in each brain hemisphere could provide the means by which a sound is judged to originate from one side or the other. Unlike with ITDs, there is no requirement for exquisite temporal sensitivity in the excitatory or inhibitory inputs to the LSO. Consequently, a number of studies have demonstrated *de novo* generation of ILD sensitivity at neural stages beyond this.

Maps of Auditory Space

The extent to which any brain region provides for an integrated representation of all of the auditory spatial cues remains to be determined. The superior colliculus (SC) of the midbrain appears to contain a map of auditory space aligned with that of the visual world. Neurons in the middle layers of the SC respond to light and sound arriving from the same direction in two-dimensional space, and project onto motor neurons in the deep layers of the SC to facilitate head- and body-orienting responses. Although such a map is clear evidence for integrating visual and auditory spatial cues, it is unclear the extent to which all of the spatial cues contribute to the space map in the SC; evidence exists to suggest that ITD, the main cue used by many mammals for localization in the horizontal plane, does not contribute to this map.

The clearest evidence for such a map in the SC again comes from studies in the barn owl—a hearing specialist that can use sound alone to catch its prey. However, these animals are sensitive to ITDs for frequencies up to 10 kHz, well beyond the upper limit observed in other species. Further, barn owls, because of the relative geometry of their ears, uniquely employ ILDs for localization in the vertical, rather than horizontal, plane. The presence of two cues, one for azimuth (the horizontal plane) and one for elevation, over the same frequency range, renders such a two-dimensional mapping of binaural cues possible. For most other species, however, the restriction of ITD sensitivity to low frequencies, and ILDs to high frequencies, means that the two binaural cues operate independently over different spectral regions. Aside from the visually controlled SC map, there is little evidence in mammals for any organization in the ascending auditory pathways that provides for orderly maps of auditory spatial cues such as ITD. Evidence for such a map in the MSO is weak at best, and the anatomical projection pattern of the MSO to the major auditory structure in the midbrain, the inferior colliculus (IC), does not appear to be organized in any fashion that would suggest a map of auditory space. This could be taken to suggest that an *orderly* map of space is not required in the auditory system to determine the source location. Maps of space in other modalities of vision and touch arise because the peripheral sensors at the back of the eye and on the surface of the skin are spatially arranged maps of visual and somatosensory space, respectively, in the same way that the peripheral sensors in the inner ear are a spatially arranged map of sound frequency.

The Neural Code for Interaural Time Difference

The form of the neural code for ITD is the subject of ongoing debate. According to Jeffress' model, ITD-sensitive neurons indicate the location of a sound source by virtue of their maximum firing rate—they are tuned to a preferred ITD. However, recent studies have called into question the basic tenets of this model. Apart from questioning the means by which individual neurons are tuned to their preferred ITD (axonal conduction delays vs. neural inhibition), experimental recordings from a range of small mammals suggest that neurons are best able to discriminate between different values of ITD where their response changes most rapidly as the ITD is shifted. This generally occurs on the slopes rather than the peaks of the firing rate functions, and possibly accounts for why so many ITD functions have their most sensitive slope around zero ITD, where auditory spatial acuity is known to be greatest, rather than at their peaks. Indeed, in a manner similar to that of the paired LSOs, the activity in each MSO may act in a "push-pull" manner, one MSO more active when the source is located to one side, and the other MSO more active when the source is located to the other. The relative activity in each MSO would therefore indicate the location of the sound source.

Cortical Codes for Auditory Space

Although the representation of auditory space in higher brain centers, such as the neocortex, remains to be determined, evidence is emerging that it takes the form of a panoramic representation; that is, neurons respond to sounds in almost all locations on the opposite side of space. The relative contributions of binaural and spectral cues to this tuning remain to be determined. Regardless of the precise nature of the underlying mechanisms, the form of the neural code has been a matter of great interest. Although source location could be encoded by spike rate, a code for auditory space based on spike timing has also been suggested, and several models have suggested that spike-timing information alone is sufficient to replicate localization performance. By whatever means the various spatial cues are combined into a single representation of auditory space, however, it is clear that by some level of cortical process an integrated representation of auditory space is accomplished, and human cortical imaging studies have started to elucidate specific brain regions potentially contributing to the various aspects of spatial sound, including source location and motion, and the spatial extent of such sources.

David McAlpine

See also Auditory Localization: Psychophysics; Auditory Processing: Central

Further Readings

Campbell, R. A., & King, A. J. (2004). Auditory neuroscience: A time for coincidence? *Current Biology, 14*(20), R886–888.

Carr, C. E., & Konishi, M. (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *Journal of Neuroscience, 10,* 3227–3246.

King, A. J. (2004). The superior colliculus. *Current Biology, 14*(9), R335–338.

McAlpine, D., & Grothe, B. (2003). Sound localization and delay lines—Do mammals fit the model? *Trends in Neurosciences, 26*(7), 347–350.

AUDITORY LOCALIZATION: PSYCHOPHYSICS

Humans and most other animals can locate the source of sound in all three spatial dimensions: left-right (horizontal or azimuth) dimension, up-down (vertical) dimension, and near-far (distance or range) dimension. The sound produced by a sound source has no spatial properties, only the physical properties of frequency, intensity, and time. Thus, the ability to locate the position of a sound source is based on cues that result from the interaction of sound and objects (e.g., the head and the body of the listener or reflective surfaces) in the path of the sound as it travels from the source to the ear canals and ear drums on each side of the head. A different set of cues is used to determine the location of a sound source in each of the three spatial dimensions. This entry describes the psychophysics of locating auditory signals for each spatial dimension and also considers lateralization (when signal is presented through headphones) and localization in reflective spaces.

Horizontal Dimension

A sound from a source that lies to one side of a listener in the azimuth dimension such as a source off to the right will arrive at the ear closer to the sound before it reaches the ear on the other side of the head. This small interaural (between ears) time difference in the arrival of the sound at the two ears produces an interaural time difference (ITD) cue for locating sound sources in the azimuth dimension. The sound will also be more intense (have a higher sound level) at the ear closer to the sound source relative to that at the opposite ear. Thus, there is also an interaural level difference (ILD) cue for sources that lie in the horizontal dimension (see Figure 1). When the sound source is directly in front of a listener, the ITD and ILD cues are both zero because the sound source is midway between the two ears. As the location of the sound source is moved to one side or the other, the ILD and ITD cues increase. Thus, large ITD or ILD values indicate sound sources that are located well to one side of the listener.

The ILD cue is based largely on the fact that the head and body shield the sound on the ear opposite the sound source relative to that on the side of the sound source. That is, the head provides a sound shadow for the sound that arrives at the ear opposite the location of the sound source. This sound shadow produces a lower level sound at the far ear. The amount of attenuation of the level of the sound at the far ear depends on the sound's frequency. Higher frequencies receive more attenuation than do lower frequencies (the short wavelengths of high-frequency sounds allow high-frequency sounds to be reflected from the head), resulting in large ILDs for high frequencies compared with low frequencies. Thus, ILDs provide a better cue for sound source location in the azimuth dimension for high frequencies relative to low frequencies. For several reasons, the ITD cue provides most accurate information about the location of a sound source in the azimuth dimension when the sound has low frequencies (below about 1,200 hertz [Hz]). Above about 1,200 Hz, sounds such as sinusoidal and other narrow-band sounds cannot be located on the basis of an ITD. Thus, ITDs provide the best cue for sound source location in the azimuth dimension for low frequencies.

Humans are good at localizing the azimuth position of a sound source at low and high frequencies but are less accurate for mid frequencies (between 1,200 and 2,000 Hz). This and other results support the duplex theory of sound localization, which contends that in locating sounds in the azimuth dimension, ITD cues are used for low-frequency sounds and ILD cues are used for high-frequency sounds. For mid-frequency sounds, neither cue is as useful and localization acuity suffers.

Humans can detect a difference in ITD as small as 10 one-millionth of a second (10 microseconds) and a difference in ILD as small as 0.5 decibels (dB). If the location of a sound source directly in front of a listener is changed by about one degree of visual angle, the listener is likely to perceive the change in location, but about seven degrees is required to detect that the location of a sound source has changed when the source was originally located to the side opposite an ear.

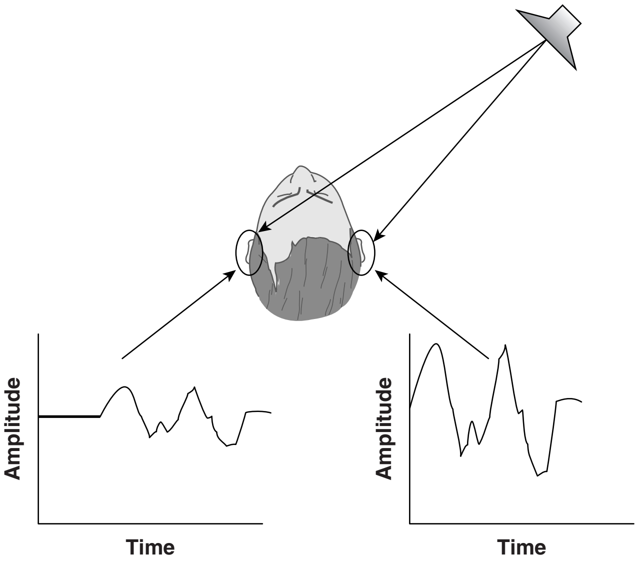


Figure 1 Interaural Differences in Arrival Time and Sound Level

Source: From Yost (2007).

Notes: A sound off to one side of a listener produces an interaural difference in arrival time (interaural time difference, ITD) and an interaural difference in sound level (interaural level difference, ILD).

Vertical Dimension

A sound source located directly in front of a listener and one located directly over head and one located directly behind a listener all are midway between the ears, resulting in all of these locations producing zero ILDs and ITDs. These sound sources are located on the mid-sagittal plane that describes the location of all positions around a listener such that the sources are midway between the ears. The mid-sagittal plane is a cone of confusion because the ILD and ITD cues are all the same value. A cone of confusion is any situation in which the ILDs and ITDs are the same for the spatial location of several different sound sources. This means that ILDs and ITDs cannot be used to determine the location of the sound source positioned on a cone of confusion. Yet, listeners can locate the position of sound sources located on cones of confusion.

The ability to locate the sources of sounds on cones of confusion means that a different set of cues than ITDs and ILDs must be used for localizing the position of sound sources in the vertical dimension and to avoid front–back confusions (confusing a sound source in front for one in back when both sources produce the same ITD and ILD). That is, some other interaction between sound and objects (e.g., the body, head, and ears) in the path of the sound must be used to provide vertical localization cues besides those producing ITDs and ILDs.

As the sound from a source travels across the body, head, and ears (especially the pinna, which are the external parts of the ears), the sound encounters many small obstacles (e.g., ridges in the pinna) that impede and delay its transmission to the ear canals. Thus, the sound arriving at the ear canal is different from that at the source because of the many parts of the body the sound encounters. The sound's spectrum, especially in the high frequencies, is altered as the sound is transferred from its source to the ear canals. A head-related transfer function (HRTF) describes these spectral changes. The levels of high-frequency sounds are especially changed by these various body parts. Exactly how the sound changes depends on the angle of the sound source relative to the ears. Thus, the characteristics of the HRTF, especially in the high frequencies (above 6,000 Hz), depend on the location of the sound source in the vertical direction, and the HRTF is different for a sound source located in front from that located in back of a listener (e.g., the pinna stick out from the side of the head blocking sound coming from a source behind a listener). Thus, features of the HRTF at high frequencies provide potential cues for locating sounds in the vertical dimensions and on cones of confusions. Figure 2 depicts the left-ear and right-ear HRTF amplitude spectra for sound sources occurring at four locations in the mid-sagittal plane. Most of the changes in these spectra as the vertical position of the sound source changes are at high frequencies. Note, for instance, that the frequency of the first noticeable valley in the spectrum (at approximately 7 kilohertz [kHz] for the lowest HRTF spectrum for the left ear) shifts upward in frequency in both ears as the position of the sound source moves up in the vertical direction.

Human listeners are not as good at localizing sounds in the vertical direction as they are at locating sounds in the horizontal direction. Sounds that contain only low frequencies are poorly localized in the vertical dimensions or along cones of confusion because, as Figure 2 indicates, there is little change in the HRTF spectra at low frequencies. Listeners have an especially difficult time separating sound sources located in front from those located behind them (front-back and back-front confusions occur frequently) when the sources are located on cones of confusion. Moving one's head can overcome such confusions, most likely because the auditory system receives several different estimates of ITD, ILD, and HRTF cues as the head moves, and it can interpret these several cues to aid in sound source localization.

Distance Dimension

The ability to locate how far a sound is from a listener is especially poor. Two cues can probably be used to determine the distance of a sound. The loudness of sound varies with distance—the greater the distance between the location of a sound source and a listener, the softer the sound. However, the actual level of a distant sound could be increased, which would increase its loudness even if the sound source's position was not changed. Thus, loudness is only a reliable cue for sound source location if the listener knows something about the expected loudness of the sound—for

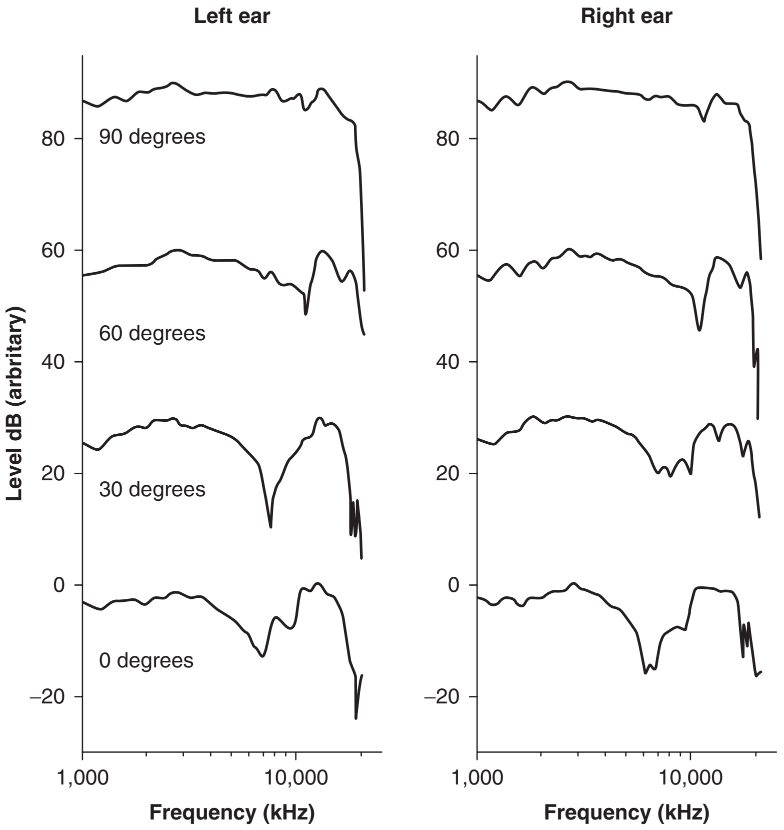


Figure 2 Left Ear and Right Ear HRTF Amplitude Spectra Measured at Four Different Elevations Relative to the Listener's Head

Source: From Yost (2007).

Notes: Note the spectral valley that is around 7 kHz for the 0 degree measurements increases in frequency as the elevation of the sound source rises. This increase in the frequency of the spectral valley may be a cue this listener can use to localize the vertical location of sound sources.

example, for a sound source that is moving away, or a familiar sound, such as a person speaking.

The other potential cue for distance results from the interaction of sound with reflective surfaces the sound may encounter. In a room, a sound source that is close to a listener will arrive at the ears of the listener with a far greater level than sounds arriving from any reflective surface (e.g., a wall) because the sound must travel some distance to the reflective surface and then from the surface back to the ears of the listener (the greater the distance a sound travels, the lower its sound level will be). But, if the sound source in this room is moved further away from the listener, the difference in the level at the listener between the sound from the source and that from the reflective surface is likely to be smaller. Therefore, the direct-to-reflective ratio of sound level is a potential cue for locating the distance of a sound source. Large direct-to-reflective ratios of sound level would be associated with close sound sources. Indeed, locating the distance of sound is better in reflective spaces than in spaces without reflective surfaces (e.g., outdoors). Thus, there are three classes of cues for locating the position of a sound source: interaural differences of time and level (ITDs and ILDs) for localization in the azimuth direction, high-frequency characteristics of the HRTF for vertical localization and that involving cones of confusions, and direct-to-reverberate sound level and, on occasion, loudness for distance.

Lateralization and Localization

The study of sound localization can involve actual sound sources (e.g., loudspeakers and headphones) and variables such as ITDs can be manipulated to determine how the auditory system process a particular cue (e.g., ITDs). Lateralization refers to the perception of sound images when the sounds are presented over headphones and localization is used to describe the perception of sounds from actual sound sources or loudspeakers. Presenting sounds over headphones provides a great deal of experimental control of the variables used in sound localization. In many situations, sounds presented over headphones appear to lie inside the listener's head and to move (e.g., left or right) based on the variable that is manipulated (e.g., an ITD = 0 places an image in the middle of the head, and an ITD = 300 microseconds places the image toward the ear that receives the sound first). Sounds presented via loudspeakers are usually perceived as real sources outside the head out in the real world. That is, sounds presented over headphones are not perceived as actually being out in the real world where actual sound sources exist. Using technology based on HRTFs, sounds presented over headphones can be perceived as if a sound source was at an actual location in three-dimensional space. Such a virtual perception of sound source location requires careful control of the stimuli delivered by headphones to each ear before a veridical perception of a real sound source in the real world can occur.

Localization in Reflective Spaces

In a reflective environment (e.g., a room), sound arrives at the listener from the source but also from the many reflective surfaces. Each reflective surface is like a different sound source. Thus, the auditory system receives sounds from many different locations when the sound source is in a reflective environment. Yet, except in extremely reflective places, listeners are able to accurately judge the location of the actual sound source and are not confused by the sound arriving from reflective surfaces. This fact is referred to as the *precedence effect,* indicating that the sound from the source takes precedence in sound source localization over that from reflective surfaces. This is partially because the sound from the source reaches the listener before any sound from reflective surfaces because the sound from a reflective surface must travel a longer path (from the source to the surface and then to the listener). Thus, the auditory system appears to suppress information about a possible sound source location from later arriving copies (the reflected sound) of the earlier arriving sound from the originating sound source.

Thus, the auditory system uses a variety of cues to localize the source of sounds in many contexts including reflective spaces.

William A. Yost

See also Acoustics and Concert Halls; Auditory Localization: Physiology

Further Readings

Blauert, J. (1997). *Spatial hearing.* Cambridge: MIT Press.

Popper, A. N., & Fay, R. R. (2005). *Sound source localization.* New York: Springer.

Yost, W. A. (2007). *Fundamentals of hearing: An introduction* (5th ed.). San Diego, CA: Academic Press.

Yost, W. A., & Dye, R. (1991). Properties of sound localization by humans. In R. Altschuler, D. Hoffman, R. Bobbin, & B. Clopton (Eds.), *Neurobiology of hearing: The central nervous system* (pp. 187–199). New York: Raven Press.

Yost, W. A., & Gourevitch, G. (Eds.). (1987). *Directional hearing.* New York: Springer.

AUDITORY MASKING

Noise is everywhere nowadays. Often, one cannot carry on a simple conversation because it is impossible to hear or to be heard over the din of competing background sounds. What hinders communication in these cases is auditory masking. By definition, *auditory masking* is the ability of one sound to make another sound hard to hear. This entry briefly describes the laboratory techniques for measuring auditory masking, the effects of major factors and their interactions, and the possible causes of auditory masking.

Measurement

In the research literature on auditory masking, the sound to be detected is referred to as the *signal,* and the interfering sound is the *masker.* Audibility of the signal in the presence of the masker is measured through many experimental trials. A typical procedure involves presenting on each trial an instance of the signal-plus-masker and masker alone in random order and having the listener judge which of the two sounds contained the signal. This procedure ensures that any bias the listener might have to overreport or underreport the presence of a signal will not influence the results. After the listener has received a period of practice in the task, the level of the signal is found that produces a criterion level of percent correct detection performance. This level is taken as the *masked threshold* for the signal. The difference in decibel (dB) between the masked threshold and the threshold for the signal in quiet is then taken as a measure of the *amount of masking*.

Effects of Factors

The types of signals and maskers that have been studied most extensively in research on auditory masking are speech, pure-tones, and bands of noise. The amount of masking produced by any one of these sounds depends on a complex interplay of factors involving the sound's intensity and its spectral, temporal, and spatial properties in relation to the signal. Years of research in audition have mapped out in great detail the major effects of these factors and their interactions.

As might be expected, increasing the intensity of the masker generally increases the amount of masking regardless of the type of signal or masker one might choose. The rate of increase, however, is generally greatest when the masker spectrum falls just below the frequency of the signal: an effect referred to as the *upward spread of masking*. If the level of the masker is fixed, the amount of masking will also decrease rapidly as the masker is moved away from the signal in frequency. The rate at which masking decreases provides a measure of the *frequency selectivity* of the auditory system, the degree to which the ear is capable of resolving individual details in the spectrum of sounds.

Sounds also have masking effects that extend over time. *Forward masking* occurs when the masker precedes the signal in time; *backward masking* occurs when the masker follows the signal in time. Little backward masking is obtained for practiced listeners unless the signal is quite brief and the masker follows the signal closely in time. Forward masking, on the other hand, is obtained for much longer duration signals and extends over a much longer period following the offset of the masker. The time course of forward masking provides one measure of auditory *temporal resolution*, the ability of the ear to follow rapid fluctuations in the amplitude of sound over time. Forward masking is also commonly used as a means to study complex processes occurring in the cochlea that are known to affect auditory frequency selectivity.

No less important to masking is the spatial dimension of sound. Everyday experience tells us that the interference among sounds is less when they originate from different locations than when they originate from the same location. This is certainly borne out by research. However, most of the research has actually been conducted with sounds played over headphones to permit precise control over the relevant spatial cues. One important spatial cue is the difference in the time of arrival of the sound at the two ears, referred to as the *interaural time difference* (ITD). Masking is generally less when the ITD of the signal is made to differ from that of the masker. The reduction in masking is referred to as the *binaural masking level difference* and can amount to as much as 15 dB. Such sizable reductions in masking attest to the importance of spatial cues for normal listening in everyday noisy environments.

Other factors affecting masking become evident when sounds are constructed in the laboratory to more closely approximate those encountered in the real world. Trial-by-trial variation in the spectra of maskers, intended to imitate the type of variation found in natural sounds, can significantly elevate masked threshold for a pure-tone signal over the course of trials, even when masker frequencies are prevented from falling anywhere near the signal. Qualitative similarities between signal and masker can have an equally profound effect, as when amplitude fluctuations of the masker covary with the signal over time or when a speech signal is masked by speech versus noise. *Informational masking* is the term often used to describe such masking effects that cannot be accounted for by known factors involving the spectral, temporal, or spatial overlap of signal and masker.

Causes

Auditory masking occurs for different reasons at different levels of the auditory system. If the signal and masker are close in frequency, they will produce overlapping patterns of vibration along the length of the basilar membrane and so compete for the response of the same auditory nerve fibers—a type of *line busy effect*. The masker may also *suppress* the response to the signal at this level because of active processes that are known to occur in the cochlea. Further along in the chain of processing, auditory nerve fibers undergo *adaptation* in response to the masker. Masking may therefore result if the signal occurs before nerve fibers have had time to recover, as in the case of forward masking. Finally, maskers may have effects on *attentional* or *perceptual processes* occurring at a more central level of the auditory system. Such effects are believed to play a role in informational masking though they are not well understood. This continues to be an active area of research.

Robert A. Lutfi

See also Auditory Frequency Analysis, Neural; Auditory Frequency Selectivity; Signal Detection Theory and Procedures

Further Readings

Kidd, G. (2007). Informational masking. In W. A. Yost & A. N. Popper (Eds.), *Springer handbook of Auditory research: Auditory perception of sound sources.* New York: Springer.

Yost, W. A. (2007). Masking. In *Fundamentals of hearing: An introduction.* New York: Elsevier.

AUDITORY PROCESSING: CENTRAL

The two main functions of hearing are localization and identification of sounds. *Central auditory processing* deals with the coding of sounds into neural representations in the central auditory pathways, but mostly at the level of the cerebral cortex. In other words, research on central auditory processing is interested in the perception, storage, and (re) cognition for purposes of behavior of different types of sounds, including communication sounds and music. How are the neural codes found and tested? Neural activity patterns in the brain in response to acoustic stimuli can be measured noninvasively (via blood flow and blood oxygenation changes) with functional imaging techniques in humans or, at much greater precision using microelectrodes, with electrophysiological techniques in animals. In addition, this entry addresses what auditory representations may have in common with other sensory systems, such as the visual system, and what may be special about them.

Tonotopic Organization

Since the days of Hermann von Helmholtz, the auditory system has been considered to function primarily as a frequency analyzer. According to work by Georg von Békésy, who was awarded the Nobel Prize in 1961, sound reaching the eardrum generates a traveling wave in the cochlea of the inner ear. Depending on the frequency of the sound, the traveling wave achieves maximum amplitude in different locations along the *basilar membrane* of the cochlea. Thus, sound frequency (from low to high) is translated into a place code, along a tonotopic axis, which is preserved from the cochlea through the processing stations of the brainstem (e.g., inferior colliculus, IC) and thalamus (medial geniculate nucleus, MGN), all the way to the auditory cortex.

Topographic organization in the auditory pathways (*tonotopy* or *cochleotopy*) can be demonstrated with electrophysiological mapping studies using tones of a single frequency (pure tones) and testing the responses of single neurons to such stimuli. Usually, a best frequency can be found that evokes the highest firing rate (or number of action potentials fired) in response to such a tone. At early stages of the central auditory pathways, best frequencies are arranged in an orderly fashion—a tonotopic map. In the primary auditory cortex of humans or monkeys, for example, neurons at its anterior end respond best to low frequencies (a few hundred hertz [Hz]), whereas neurons at the posterior end fire most to high frequencies (10 kilohertz [kHz] or above). Tonotopic organization is analogous to retinotopy in the visual system and somatotopy in the somatosensory system, in the sense that stimuli close to each other along a physical domain are mapped into positions next to each other in the brain. Topographic maps provide for efficiency of coding and shorter wiring of local neural circuits, and they facilitate the orderly development of sensory systems.

Processing of Complex Sounds

One problem that the frequency analyzer theory has not solved is how information from different frequency channels gets integrated, or how an organism analyzes complex sounds, that is, sounds that vary in both frequency and time. Most natural sounds are complex sounds, so solving this problem is of the essence in the analysis of higher auditory pathways. The problem has been tackled successfully in a number of specialized systems, such as frogs, songbirds, barn owls, and echolocating bats. For all these species, a neuroethological approach has been adopted based on functional-behavioral data and use of complex natural sounds for stimulation. The same approach has been used more sparingly in higher mammals, including primates.

The neurophysiological basis in humans for processing complex sounds, such as speech or music, can rarely be studied directly with invasive methods at high resolution, except sometimes before epilepsy surgery. Therefore, animal models, such as cats or nonhuman primates, have to be used. The question then arises to what extent it is valid to apply human speech sounds as stimuli for the study of neurons in animals, when we know that their brains did not evolve to process human speech but communication sounds from their own species. From a biological-evolutionary vantage point, it seems more meaningful, therefore, to employ species-specific vocalizations. In this case, we can be confident that the central auditory system of the studied species must be capable of processing these calls.

When comparing human speech sounds with communication sounds in other species we can see, however, that most systems have certain components in common, that are used as carriers of semantic (meaning-related) information. Among them are brief segments of frequency changing over time (e.g., formant transitions in human speech) or noise bursts with specific center frequencies and bandwidths. Such universal elements, alone or in combination, make up phonemes, that is, the basic units of human speech, but are also common in animal communication sounds. Thus, it is meaningful to use them as stimuli for probing central auditory structures that are intermediate between the primary auditory cortex (where neurons respond best to tones) and what would be considered the language cortex in humans (where more higher-level processing of complex stimuli is carried out). Similarly, to study how the brain processes music, we can start with the use of complex harmonic tones or chords. In the case of both speech and music processing, we would also need to consider how the short constituent elements (syllables, words, and chords) are tied together, at the neural level, to form longer sequences of sounds, such as sentences and melodies. Auditory sequence processing may require the use of the same or similar structures as the programming of motor sequences, including the premotor cortex, basal ganglia, and cerebellum.

Hierarchical Organization and Processing Streams

Neurons selective to the early elements of speech and music have indeed been identified in various species at the next stage of processing, which follows the primary (or *core*) auditory cortex and is often referred to as the auditory *belt* region. The belt surrounds the core areas and, together with the next stage, the *parabelt,* forms the totality of the auditory cortex proper in higher mammals with more than a dozen discrete auditory cortical areas (Figure 1, next page; see legend for abbreviations. See also color insert, Figure 2). In addition, auditory-responsive regions exist in the anterior superior temporal (aST) cortex (extending laterally into the superior temporal sulcus and beyond), premotor cortex, prefrontal cortex (ventrolateral [VLPFC], dorsolateral [DLPFC]), inferior parietal cortex (IPL), and even in the cerebellum (not shown).

Thus, many more auditory cortical areas exist than previously thought, and they are arranged in a processing hierarchy where discrete steps occur at every level. Auditory cortical areas form at least two major processing streams, an anterior-ventral one for the processing of *what* a sound represents, and a posterior-dorsal one for *where* in space a sound source is located (Figure 1; see also color insert, Figure 2). The two main functions of hearing, as defined earlier, therefore seem to be realized in two largely segregated cortical processing streams, similar to what happens in the visual system. Although the role of the ventral stream in object identification is virtually undisputed in either system, some researchers have sought to expand the role of the dorsal stream beyond spatial processing and have postulated its involvement more generally in sensory-guided action (i.e., *how* we perform certain acts).

Hierarchical models sometimes come across as simple feedforward models, neglecting the rich descending pathways at every level, from the cortex back to thalamus and brainstem, and the reciprocity of cortico-cortical connections. However, the existence of back-projections is often inherent in arguments made about *top-down processing,* whereby hierarchically higher structures by virtue of descending pathways influence activity in lower-order

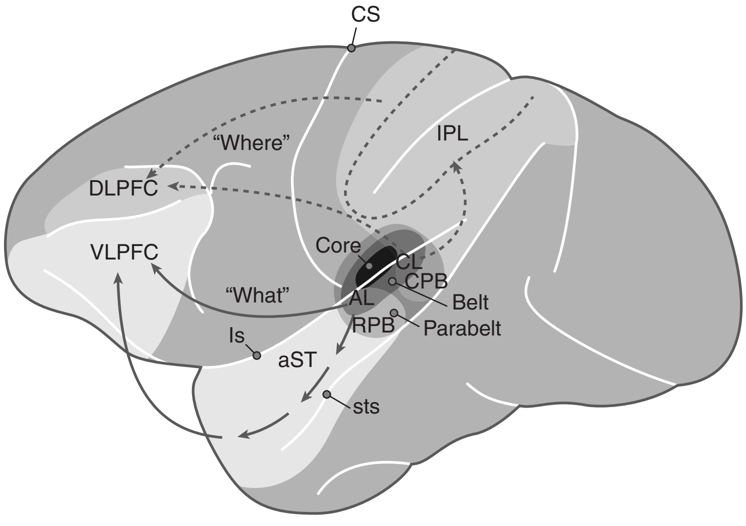


Figure 1 Schematic Illustration of Hierarchical Processing in the Auditory Cortex of the Rhesus Monkey and Processing Streams for "What" and "Where"

Notes: Auditory signals arrive first in the core areas (consisting of the primary auditory cortex, A1, and two rostral areas) from subcortical inputs in thalamus and brainstem. Activity then propagates to the belt (AL, CL, among others) and parabelt areas (rostral and caudal, RPB and CPB, respectively), which give rise to two pathways projecting to two largely segregated regions in the prefrontal cortex (PFC): the ventrolateral and dorsolateral (VLPFC and DLPFC) regions, respectively. The ventral processing stream is also relayed through the anterior superior temporal (aST) cortex, where regions or patches specialized for the processing of voices and communication sounds have been found in both monkeys and humans. The dorsal stream, in addition to its projection to the DLPFC, is relayed also through the inferior parietal lobule (IPL) of the posterior parietal cortex (PPC, particularly the ventral intraparietal region, VIP). Rich back-projections exist from the prefrontal cortex to the PPC as well as to the aST (not shown). See also color insert, Figure 2.

stations. Descending pathways also are able to adjust the response properties of neurons at lower levels and therefore mediate many forms of auditory plasticity.

Auditory Object Identification

This entry now discusses the processing hierarchy in the anterior-ventral auditory "what"-stream. Most notably, in the lateral belt areas of the superior temporal (ST) region in rhesus monkeys, a large proportion of neurons selective to brief frequency changes as well as neurons tuned to certain bandwidths of noise or combinations of tones have been found. When tested with whole monkey vocalizations, more neurons are found in the anterior than in the posterior part of the lateral belt that respond selectively to these calls. In the anterior parabelt and adjacent ST regions of rhesus monkeys, even more neurons respond well and quite selectively to monkey vocalizations. A region anterior and lateral to the primary auditory cortex also contains neurons that respond to the pitch of a complex tone even when the fundamental frequency is removed. The existence of such "pitch neurons" is important because pitch is critical for identifying and segregating auditory objects, especially in the context of music and speech.

The question arises by what neural mechanisms such selectivity for complex sounds is generated. Studies in which monkey calls are dissected into their constituent elements (both in the spectral and temporal domains), and the elements are then played to the neurons separately or in combination can provide an answer to this question. A sizable proportion of neurons in the anterior lateral belt respond much better to the whole call than to any of the elements. These results are indicative of nonlinear summation (*combination sensitivity*) in the frequency and time domain playing a crucial role in the generation of selectivity for specific types of calls. Similar mechanisms had previously been described in bats and songbirds. Coincidence detection in the time domain is perhaps the most important mechanism in shaping this selectivity. Temporal integration, the time window during which inputs are summed, extends 100 to 200 milliseconds, which fits well with the fact that most "syllables" in monkey calls (as well as in human speech) are of that duration.

Neuroanatomical studies have shown that anterior lateral belt projects to the VLPFC, which also receives visual projections and has a function in object working memory. Categorization of communication sounds and the solution of the invariance problem (i.e., the fact that we can recognize a particular speech sound regardless of pitch or accent of the speaker's voice) may be performed with the help of VLPFC. Thus, the lateral belt areas are not the ultimate stage in the processing of communication sounds. They only present an intermediate stage, similar to areas in the visual system that contain neurons selective for the size of visual stimuli. Visual cortical neurons selective for more specific objects, such as faces, are not found until even higher processing stages. In the auditory cortex, areas in the anterior ST regions (as part of the ventral processing stream) are the ones that seem to specialize in the processing of auditory objects, including voices and vocalizations. Functional magnetic resonance imaging (fMRI) studies in monkeys have confirmed the existence of a prominent patch in the anterior ST region that is activated selectively by species-specific communication calls. In humans, the anterior ST region (not its posterior part, as had previously been thought) also contains a processing hierarchy for phoneme identification as well as a "voice center" for the identification of human voices.

Auditory Space Processing

The second main task of hearing is to localize sound sources in space. The auditory periphery (the basilar membrane with its inner and outer hair cells), however, cannot a priori encode a two-dimensional (2-D) spatial signal, as do the visual and somatosensory peripheries (retina and skin surface, respectively), because it extends only along a single dimension. Auditory space, therefore, has to be computed from attributes of sound that vary systematically with spatial location and are thus processed differentially by the central auditory system. This problem is logistically similar to the computation of three-dimensional (3-D) spatial information from 2-D sensory information in the visual system, where the 2-D signals from each retina are combined to render 3-D perception (stereoscopic vision). Sound attributes most commonly assigned to spatial quality are differences between sound arriving at the two ears. Both the intensity (level) and the time of arrival of sound originating from the same source differ when the sound source is located outside the median plane, such as to the left or right. A sound to the right reaches first the right ear, then the left (and vice versa), which leads to an interaural time difference (ITD). Sounds are also less intense on the other side of the head, which leads to an interaural level difference (ILD). ITD and ILD are registered and mapped already in areas of the brainstem, such as the superior olivary complex, and are passed along to higher centers, including the cortex, from there. In addition to these binaural cues, monaural spectral information contributes to auditory space perception: The frequency composition of a sound arriving at the eardrum from different directions in space varies because the external ears (pinnae) and the head have certain spectral filter characteristics. In other words, sound bouncing around in the pinnae leads to the amplification of certain frequencies and the attenuation of others. Specific spectral "fingerprints" can be assigned to different spatial locations, with attenuation of particular frequency bands ("spectral notches") varying systematically with azimuth (left/right) or elevation (up/down). Spectral cues are the most important cues for determining the elevation of a sound and for resolving front/back confusions, where binaural cues would be identical.

The information computed by lower brainstem structures is used by higher centers of the midbrain, such as the inferior and superior colliculi, to guide orienting movements toward sounds. For more "conscious" spatial perception in higher mammals, including humans, however, the auditory cortex is indispensable, as cortical lesions almost completely abolish the ability to judge the direction of sound in space. Neurons in the primary auditory cortex show tuning to the spatial location of a sound presented in free field. Many more neurons that are sharply tuned to spatial location are found in the posterior (or caudal) regions of lateral belt and parabelt of monkeys. Lesions in these regions cause a marked deficit in sound localization, whereas lesions in anterior parts of the belt cause auditory pattern recognition deficits. Functional neuroimaging studies in humans also demonstrate specific activation in the posterior auditory belt and parabelt (leading into the inferior parietal cortex) while subjects localize sounds in space or listen to sounds moving in space.

Neuroanatomical studies confirm that posterior lateral belt gives rise to a processing stream that is largely segregated from the one originating in anterior lateral belt: the posterior lateral belt projects to the DLPFC, which is known to be important for spatial working memory.

Dual Processing

Both animal and human studies suggest that information about auditory patterns or objects is processed, among others, in the aST region. By contrast, auditory spatial information is processed mainly in posterior ST and parietal regions of the cortex. Both pathways continue into the PFC, where they innervate largely segregated target regions in the ventrolateral and dorsolateral PFC, respectively. This dual processing scheme is reminiscent of the visual pathways, where a ventral stream has been postulated for the processing of visual object information and a dorsal stream for the processing of visual space and motion. Of course, in neither system is the separation between functional streams complete. Whether one thinks of such "crosstalk" as technical imperfection or necessary convergence of information is a matter of interpretation. There are also alternative accounts of this functional segregation in both sensory systems, which argues more for perception than action. In this view, the role of the dorsal stream would be extended from purely spatial processing to providing feedback in auditory-motor transformations.

Josef P. Rauschecker

See also Auditory Processing: Peripheral; Auditory System: Structure; Brain Imaging; Cortical Organization; Modularity; Music Cognition and Perception; Neural Recording; Neural Representation/Coding; Physiological Approach; Speech Perception

Further Readings

Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature, 436,* 1161–1165.

Blauert, J., (1996). *Spatial hearing* (2nd ed.). Cambridge: MIT Press.

Helmholtz, H. von (1954). *On the sensation of tones.* New York: Dover. (Original published 1885)

Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences USA, 97,* 11793–11799.

Knudsen, E. I., & Konishi, M. (1978). A neural map of auditory space in the owl. *Science, 200*(4343), 795–797.

Merzenich, M. M., & Brugge, J. F. (1973). Representation of the cochlear partition on the superior temporal plane of the macaque monkey. *Brain Research, 50,* 275–296.

Petkov, C. I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., & Logothetis, N. K. (2008). A voice region in the monkey brain. *Nature Neuroscience, 11,* 367–374.

Rauschecker, J. P., & Scott, S. (in press, 2009). Maps and streams in the auditory cortex: How work in non-human primates has contributed to our understanding of human speech processing. *Nature Neuroscience.*

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences USA, 97,* 11800–11806.

Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science, 268,* 111–114.

Recanzone, G. H., & Sutter, M. L. (2008). The biological basis of audition. *Annual review of psychology, 59,* 119–142.

Suga, N. (1992). Philosophy and stimulus design for neuroethology of complex-sound processing. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 336,* 423–428.

Tian, B., Reser, D. Durham, A., Kustov, A., & Rauschecker, J. P. (2001). Functional specialization in rhesus monkey auditory cortex. *Science, 292,* 290–293.

Von Békésy, G. (1960). *Experiments in hearing.* New York: McGraw-Hill.

Winer, J., & Schreiner, C. (2009). *The auditory cortex: Fundamental neuroscience.* New York: Springer.

AUDITORY PROCESSING: PERIPHERAL

The auditory periphery—the outer ear, middle ear, and inner ear—is designed to collect and amplify sound waves and to separate their different frequencies into nerve pathways that can allow the brain to determine the meaning of a sound (see color insert, Figure 3). This conduction pathway is crucial to surviving in a rich auditory environment. This entry first describes the process of sound conduction and then discusses the elements of the auditory periphery.

Sound Conduction

The vibrations caused by sound stimuli are converted by the auditory system into electrical signals that are sent along the auditory nerve to the brain. The outer ear and middle ear ensure that the vibrations are transmitted efficiently to the inner ear where fluid movements occur that can be detected by the sensory apparatus in a way that means that different frequencies and amplitudes can be distinguished; these two sensations correlate with the pitch and the loudness of the sound. This process relies on a sensory structure called the organ of Corti that rests on the basilar membrane. (The cochlea is a spiral-shaped structure shaped like a snail-shell, hence its name, which is taken from the Greek word for "snail"). The basilar membrane vibrates in response to the fluid movements set up in the cochlea by sound.

The final conversion of these vibrations into electrical signals is performed by sensory cells called hair cells that are embedded in the organ of Corti. At various stages along this pathway, the vibrations are selectively amplified giving rise to fine frequency discrimination and to a wide sensitivity range.

The Outer Ear

The protrusions at the sides of our heads that most of us think of as our ears are called pinnae; these structures are the first stage of the auditory pathway, the outer ears. An inner funnel-like entrance leads to the external auditory canal. Some animals (e.g., cats) can move their outer ears to direct them toward a sound source more effectively. Most humans cannot move their outer ears but nevertheless the reflection of sounds caused by their shape leads to significant information being sent to the brain about where a sound is located in space, especially about whether it is coming from in front of or behind or one or the other side of the head in the horizontal plane.

The tapered shape of the auditory canal, the tube that conducts sound into the head (the outer side is wider than the inner side) slightly increases sound transmission to the eardrum (tympanic membrane) and, in different animals, the length and cross-sectional areas of this external auditory canal favor different frequencies, that is, sensitivity to low, middle or high notes. In humans, this favored frequency is near that used in speech, 1 to 2 kilohertz (hertz are the number of sound wave cycles that occur per second).

The Middle Ear

The *tympanic membrane,* a translucent white (when healthy) membrane, is the boundary between the outer ear canal and the middle ear. On its inner side, the tympanic membrane is connected to one of a chain of small bones called *ossicles* that transmit the vibrations set up by sound through the middle ear to the inner ear. The ossicles connect the tympanic membrane to another membrane called the oval window membrane that vibrates the fluids in the inner ear. The difference in the size of the larger tympanic membrane to the oval window membrane results in an amplification of sounds by approximately 17 times. The tympanic membrane is susceptible to injury by high sound levels (e.g., explosions) but can recover from lower levels of damage.

The three ossicles are called (from the outermost to the innermost) the *malleus* (Latin for "hammer"), the *incus* ("anvil"), and the *stapes* ("stirrup") because of their respective shapes. They are attached together by cartilage, and muscles—the tensor tympani and the stapedius—suspend the ossicles from the skull lining of the middle ear. The muscles have separate nerve connections with the brain, but their simultaneous activation can contract them and lead to "locking" of the ossicular chain. This middle ear reflex helps to prevent loud sounds from causing damage to the inner ear.

The middle ear is otherwise an air-filled cavity that is kept at the same air pressure levels outside the skull by means of the Eustachian tube, which runs from the middle ear to the mouth. If this tube is blocked by saliva, "popping" of the eardrum (rapid movements) can be detected if the external air pressure changes such as during the ascent or descent of aircraft. These are not normally dangerous but sometimes can be alarming. However, if bacteria enter the middle ear by means of this route, infections can occur that lead to a build-up of pus that prevents the ossicular chain of bones from moving. This condition is termed *glue ear* and is a common cause of hearing impairment in small children. It is treated by allowing air back into the middle ear by the insertion of grommets (small ring-like devices) that hold open gaps made surgically in the eardrum. As the bacteria are anaerobic, this influx of fresh air helps to destroy them. The grommets can then be removed and because the holes that have been made are small and carefully placed, the eardrum can heal over when they are removed.

Another problem that can occur in the middle ear is ossification (bone overgrowth) of the ossicles, primarily with ageing. The ossicles can sometimes be replaced with artificial, often ceramic, substitutes that can help to restore the transmission of the vibrations set up by sound in the middle ear.

The Cochlea

The cochlea is a spirally fluid-filled tube; it is coiled to save space in the skull. When sound vibrations result in movements of the ossicles in the middle ear, the vibrations are transmitted to the cochlear fluids by these bones pressing on a membrane stretched across an opening in the bone at cochlear base called the oval window. The fluid movements in the cochlea cause vibration of the basilar membrane, which runs along the length of the cochlear tube.

The basilar membrane changes in its physical characteristics along the cochlea. Researchers have known since the 1930s that the basilar membrane's mechanics result in the conversion of the sound frequency to coding as a position of excitation along its length, a process that results in a tonotopic (sound-to-place) map. The main discovery was made by Georg von Békésy in the 1930s. He observed that the stiffness of the basilar membrane was greater toward the basal (middle ear) end of the cochlea than at its apex. The stiffness of the basilar membrane is graded and, as a result, different sound frequencies, although they set up a "traveling wave" along the basilar membrane, cause maximal vibrations at different points along the basilar membrane. Low frequencies vibrate the membrane most at the apex of the cochlea and high frequencies most at the base, with intermediate frequencies being represented systematically along its length.

Another membrane, Reissner's membrane, stretches from the center of the cochlea to the side of the cochlea attaching to its outer wall above the basilar membrane. Thus, the cochlea has three fluid filled compartments, the central one of which is filled with a different fluid from the other two. This central fluid is called *endolymph;* it is rich in potassium ions and importantly contains calcium ions. The other two compartments, the scala vestibuli and scala tympani, contain a fluid called *perilymph.* The principal ions in perilymph are sodium ions, and the difference in the ionic composition of the endolymph and perilymph sets up an electrical gradient that is necessary to the operation of the sensory cells of the cochlea. These cells are embedded in a sensory structure called the organ of Corti. They are called *hair cells* because protruding from their apical surface are bundles of tiny protrusions called *stereocilia* that early observers thought were hairs but that are actually modified microvilli. When the basilar membrane is deflected by fluid waves in the cochlea, deflection of the stereociliary bundles produces electrical signals in the body of the hair cell to which they are attached. This is because the shape of the bundles results in the opening of ion channels in the tips of the shorter rows of stereocilia. Potassium and calcium ions flow in rapidly depolarize the hair cell (making the electrical charge inside the cell more positive). This electrical change causes the hair cell to increase the rate of firing of the nerve fibers from the auditory nerve that are connected to its base. This process is known as mechanoelectrical transduction because a mechanical signal (sound vibration) is converted into electrical signals that can be sent to the brain. This is basically a rapid switch mechanism by which the nervous system turns a mechanical stimulus into a nerve signal to the brain. The calcium ions close the mechanoelectrical transduction channels as they enter the stereocilia, a mechanism that enables the hair bundles' sensitivity to be reset via a mechanism called adaptation. This process preserves the response of the hair bundle to deflection even if they are already deflected. A second slower mechanism also occurs that regulates the tension being placed on the mechanoelectrical transduction channels by moving the upper attachment point of the tip link along the side of the stereocilium. It has been suggested that this motion is driven by the climbing and slipping of a protein called myosin (which is also involved in muscle contractions) along the actin core of the stereocilium.

There are four rows of hair cells. One row is nearest to the central spire of bone in the cochlea, the modiolus; these are called the inner hair cells. Three rows lie toward the outer edge of the organ of Corti, the outer hair cells. These cells have different functions. The inner hair cells have large numbers (16–20) of nerve fibers attached to their bases that conduct electrical signals along the auditory nerve to the brain. The outer hair cells have fewer of these afferent nerve endings (terminals that send information to the brain) and instead receive signals from the brain that affect their activity from the efferent nerve terminals. The inner hair cells are principally responsible for conveying information about characteristics of the sound that is conducted from the cochlea to the brain. The outer hair cells boost this signal in a frequency specific way by rapidly changing their length; they exert contractile and extensible responses into the basilar membrane at the frequencies being transmitted in the sound. This mechanism only works in living animals because the outer hair cells need to be physiologically active to produce these movements. This process is known as the *cochlear amplifier* because the muscle-like responses of the outer hair cells increase the movement of the basilar membrane at appropriate points along its length, thus boosting the signal at the correct frequency position and decreasing the response at the remaining points and increasing the basilar membrane's ability to separate different frequencies. The outer hair cells thus amplify the sound signals in the inner ear for whichever frequencies they are detecting, an important function because most environmental sounds are complex; that is, they consist of many frequencies that determine their meaning in the brain.

Carole M. Hackney

See also Auditory Receptors and Transduction; Auditory System: Damage Due to Overstimulation; Auditory System: Structure; Sound Stimulus

Further Readings

Ashmore, J. (2002). The mechanics of hearing. In D. Roberts (Ed.), *Signals and perception: The fundamentals of human sensation* (pp. 3–15). Basingstoke, UK: Palgrave Macmillan.

Baguley, D., & McFerran, D. (2002). Hearing impairments: Causes, effects and rehabilitation. In D. Roberts (Ed.), *Signals and perception: The fundamentals of human sensation* (pp. 65–76). Basingstoke, UK: Palgrave Macmillan.

Dallos, P. (1992). The active cochlea. *Journal of Neuroscience, 12,* 4575–4585.

Von Békésy, G. (1960). *Experiments in hearing.* New York: McGraw-Hill.

AUDITORY RECEPTORS AND TRANSDUCTION

Hearing begins with the transduction of sound energy into electrical signals in the inner ear, through a sequence of events so fast and sensitive as to be limited only by the thermal noise of molecules moving in inner ear liquids. This entry reviews that sequence with particular focus on the remarkable properties of auditory receptor cells.

Sound originates in the motion of vibrating objects, such as the vocal cords during speech. The vibration energy—a form of mechanical energy—is transmitted to air molecules and thence to the eardrum. Vibrations of the eardrum in turn move the bones of the middle ear, which push in and out, piston-like, on the liquid-filled chambers of the hearing organ in the inner ear. The mammalian hearing organ, called the cochlea, is coiled in a spiral of several loops. Membranes running the length of the cochlea subdivide it into three tubes, or chambers, filled with liquid. The bone enclosing the upper and middle chambers has membranous, elastic windows in it. Middle ear bones push in and out on the upper chamber's window, creating pressure changes that deflect the elastic middle chamber and distend the window on the lower chamber.

The middle chamber (see color insert, Figure 3a) houses a long sensory epithelium (the organ of Corti) containing thousands of auditory receptor cells, called hair cells for their most striking feature: a "hair" bundle of fine, interlinked columnar structures. The hair cells form four long rows on top of the bottom (basilar) elastic membrane of the middle chamber. As illustrated in Figure 3(a) in the color insert, sound-driven motion of the basilar membrane bends the hair bundles against a delicate tectorial membrane. Viewed along the length of the cochlea, the motion of the basilar membrane takes the form of a traveling wave, which peaks at different places for different sound frequencies. The hair cells convert the mechanical energy of hair-bundle bending into electrical signals called receptor potentials. The electrical signals are transmitted from the inner ear to the brain via auditory nerve fibers (color insert, Figure 3a).

The mammalian cochlea has a single row of inner hair cells and three or more rows of outer hair cells (color insert, Figure 3a). Despite the small number of inner hair cells, it is their sound-evoked signals that are conveyed to brain neurons. Inner hair cells transmit signals to auditory nerve fibers across specialized points of close contact called synapses (color insert, Figure 3b), stimulating the brief electrical events (action potentials) that neurons commonly use to transfer information over long nerve fibers. The sound-evoked action potentials travel along auditory nerve fibers to neurons in hearing centers in the brain.

Outer hair cells are unique to mammals. Surprisingly, given their large numbers, they have weak contacts with auditory nerve fibers. Nevertheless, they are essential to good hearing. They amplify sound-evoked motions within the cochlea by a remarkable mechanism, electromotility, which refers to the fact that the outer hair cells change their length in response to movement of the cilia. The resulting amplification of the motion of the basilar membrane is considered critical for the unusual high-frequency capability of the mammalian ear. Other vertebrates can hear sound frequencies ranging from about 0.1 kilohertz (kHz, thousands of cycles per second) to as high as 10 kHz in some birds. Children can hear frequencies as high as 20 kHz (until hearing damage sets in), and other mammals can hear up to 50 or even 100 kHz. The entire frequency range is represented smoothly along the cochlear spiral such that the highest frequencies strongly activate hair cells at the base, next to the middle ear, and the lowest frequencies strongly activate hair cells at the top of the spiral. This systematic variation in the most effective sound frequency (best frequency or characteristic frequency) with cochlear location arises from smooth variation in the properties of the basilar membrane and hair cells.

Outer hair cells are important targets of negative feedback from the brain to the ear. Efferent signals originate in brain neurons and travel along different nerve fibers from those carrying afferent signals from the ear to the brain. Inside the cochlea, efferent signals are transmitted across synapses to afferent nerve fibers and to outer hair cells (color insert, Figure 3a). By these connections, brain neurons can reduce the electrical activity of auditory nerve fibers or the amplifying effect of outer hair cells on afferent signals.

To recapitulate, hearing begins with translation of the mechanical energy of sound into hair-bundle deflections. The hair cells transduce the deflections into electrical signals, which are amplified by outer hair cells, modulated by efferent feedback, transmitted to auditory nerve fibers, and conducted to the brain. Hearing requires further processing of the incoming signals by specialized centers in the brain; nevertheless, many important properties of hearing are determined or shaped by the inner ear. To appreciate this, we first consider in more detail how hair cells transduce the energy of sound into receptor potentials and transmit these signals to auditory nerve fibers.

Mechanoelectrical Transduction and Synaptic Transmission by Inner Hair Cells

The initial stages of transduction take place within the stereocilia, the columnar components of the hair bundle. Stereocilia are arrayed in rows of graded heights between about 1 and 10 micrometers (color insert, Figure 3c). Deflections of the hair bundle toward the tallest row are excitatory (positive); opposing deflections are inhibitory (negative); orthogonal deflections have no effect. Sound deflects the hair bundle back and forth along the preferred axis, stimulating a change in the electrical current flowing into the hair cell. This transduction current in turn changes the voltage across the hair cell's membrane, producing the receptor potential.

How does deflecting a hair bundle evoke a transduction current? The current is carried by positively charged potassium and calcium ions from the liquid bathing the hair bundle (color insert, Figure 3c). The ions enter the stereocilia through pores in special proteins (transduction channels) in the stereocilium's surface membrane. Even when the bundle is at rest, the transduction channels are open some of the time, allowing a small inward current. Deflecting the bundle positively further opens the channels, increasing the inward current. According to a widely cited model of hair cell transduction, positive deflections stretch hypothetical elastic structures, referred to as gating springs, which pull the transduction channels open. Negative deflection relaxes the gating springs, allowing the channels to close and so decreasing the inward current. Figure 3(c) in the color insert illustrates this idea in cartoon images by showing springy structures (gating springs) attached to trap doors (gates) on the external surfaces of the transduction channels. Two channels are shown closed and two are shown open with potassium ions moving through them.

The increased influx of positive ions during an excitatory stimulus creates a positive receptor potential across the hair cell membrane. The receptor potential secondarily affects the flow of current across the hair cell membrane through channels that open in response to voltage change rather than mechanical energy. These additional currents have several important effects on the afferent signal. They affect the time course of the receptor potential. The extra open channels also reduce the electrical resistance of the hair cell membrane, expanding the frequency range over which the receptor potential can follow the stimulus. One class of voltage-gated channels is critical to the transmission of the sound-evoked signal across the synapse. Voltage-gated channels that selectively pass calcium (color insert, Figure 3c) translate the receptor potential into calcium changes inside the hair cells. During a positive receptor potential, the calcium channels open and entering calcium ions activate the secretion of transmitter molecules (glutamate) from the hair cell onto the afferent nerve ending. Similar events occur at excitatory synapses throughout the brain, although transmission from hair cell to afferent neuron is remarkable for its high speed and volume. The secreted glutamate molecules cross the synaptic gap to bind to another type of ion channel, glutamate receptor channels, in the membrane of the auditory nerve fiber. The channels open, admitting positive ions into the nerve fiber; the positive voltage change then triggers action potentials. Thus, a cascade of steps involving mechano-sensitive, voltage-sensitive, and glutamate-sensitive ion channels converts the deflection of the hair bundle into the common neural code: a change in the number of action potentials per second.

Experiments that isolate components of this cascade have characterized properties, such as sensitivity and speed, that set limits on our hearing capabilities. Hair cell transduction is studied by viewing the hair bundle through a microscope and recording transduction currents with a miniature electrode while deflecting the hair bundle with a fine glass probe. Hair cells produce measurable responses to angular deflections of less than one degree. This extraordinary sensitivity allows us to detect sound intensities that move the basilar membrane by less than 1 nanometer (nm). Indeed, thermal (Brownian) motions of the cochlear liquids are the limiting noise within the system. At mammalian body temperature, transduction channels open within microseconds (10^{-6} s) of the start of a deflection, allowing such hearing specialists as sonar-using bats and certain whales to hear up to 60 to 100 kHz. The short latency is possible because of the directness and efficiency of the mechanoelectrical transduction process; the multistage transduction mechanisms of photoreceptors and olfactory neurons take much longer.

Positive hair-bundle deflections elicit transduction currents as much as tenfold larger than negative deflections of the same size. Surprisingly, this asymmetry is critical for our ability to hear at high sound frequencies. This can be seen by considering the receptor potential response to pure tones, which have a sinusoidal waveform that deflects the hair bundle back and forth symmetrically. For low-frequency tones, the transduction current and receptor potential oscillate at the same frequency as the tone. As tone frequency increases above about 1 kHz, the oscillations of the transduction current become too fast for the receptor potential to follow, so that the voltage oscillations get smaller and smaller. Because of this effect, hearing would fade away with increasing frequency above several kHz if transduction current were symmetric. Instead, the asymmetric transduction current is integrated by the membrane to produce a steady positive receptor potential, allowing us to hear at frequencies that are too fast for our receptor potentials to follow in time.

When the hair bundle is rapidly deflected to a new position and held there, transduction current activates within microseconds but then decays (adapts) with a time course ranging from tens of microseconds to tens of milliseconds. Sustained bundle deflections do not occur in the cochlea but low-frequency sounds do move the bundle slowly enough to trigger significant adaptation. Adaptation is a ubiquitous feature in sensory systems, reflecting the evolutionary advantage in emphasizing novel stimuli over background stimuli. Calcium ions that enter open transduction channels as part of the transduction current (color insert, Figure 3c) activate at least two processes that tend to close the channels. In addition to reducing the transduction current (hence, adaptation), closing the channels affects the overall stiffness of the hair bundle and therefore how much it responds to a new stimulus. This effect is most pronounced for stimuli with a similar time course as the calcium-mediated channel closure. In this way, adaptation processes help tune the hair cell to stimuli with a certain time course—that is, in a specific frequency range.

Amplification by Outer Hair Cells

Outer hair cells enhance cochlear sensitivity via a unique process that transduces receptor potentials into motions of the hair cell body. The underlying mechanism may be a voltage-induced change in the size of prestin molecules in the outer hair cell membrane. The moving outer hair cells exert force on the basilar membrane, amplifying the relative motion between basilar and tectorial membranes and therefore the deflection of the bundles of inner hair cells. The summed vibrational effect is so large that it can be detected as a sound in the outer ear canal. Measurement of these "otoacoustic emissions" is now used to screen infants for hearing loss originating in the cochlea.

Ruth Anne Eatock

See also Auditory Frequency Selectivity; Auditory Processing: Peripheral; Auditory System: Damage Due to Overstimulation; Auditory System: Evolution of; Auditory System: Structure; Auditory Thresholds; Otoacoustic Emissions

Further Readings

Fain, G. L. (2003). Hair cells and the detection of movement and sound. In *Sensory transduction* (pp. 119–158). Sunderland, MA: Sinauer Associates.

Geisler, C. D. (1998). *From sound to synapse* (pp. 91–175). Oxford, UK: Oxford University Press.

Glowatzki E., Grant, L., & Fuchs, P. (2008). Hair cell afferent synapses. *Current Opinion in Neurobiology, 18,* 389–395.

Hudspeth, A. J. (2000). Hearing. In E. Kandel, T. Schwartz, & T. Jessell (Eds.), *Principles of neural science* (pp. 590–603). New York: McGraw-Hill.

Hudspeth, A. J. (2000). Sensory transduction in the ear. In E. Kandel, T. Schwartz, & T. Jessell (Eds.), *Principles of neural science* (pp. 614–624). New York: McGraw-Hill.

Hudspeth, A. J. (2008). Making an effort to listen: Mechanical amplification in the ear. *Neuron, 59,* 530–545.

Pickles, J. O. (2008). The cochlea. In *An introduction to the physiology of hearing* (3rd ed., pp. 25–72). Bingley, UK: Emerald.

Vollrath, M. A., Kwan, K. Y., & Corey, D. P. (2007). The micromachinery of mechanotransduction in hair cells. *Annual Review of Neuroscience, 30,* 339–365.

AUDITORY SCENE ANALYSIS

When we listen, we effortlessly hear the separate sounds of our environment—voices, musical instruments, cars—each coming from its appropriate direction and having its own characteristic qualities such as pitch and timbre. This simple perceptual experience is the result of the complex brain mechanisms of auditory scene analysis (ASA), described in this entry. ASA addresses the problem of grouping together those frequency components that have originated from the same sound source, thereby separating them from other sounds that happen to be present at the same time. The problem of ASA can be conceptually divided into two parts: deciding which simultaneous frequency components belong together to form a sound source, and then sequentially tracking a particular sound source across time. These mechanisms make it possible for us to function in an auditorily cluttered environment by perceptually separating sounds that are physically mixed together in the sound waves that enter our ears. The limits of our ability to separate different sound sources are exploited in music, and our inability to fully understand the processes involved impairs the automatic recognition of speech by computer, and the effectiveness of hearing aids.

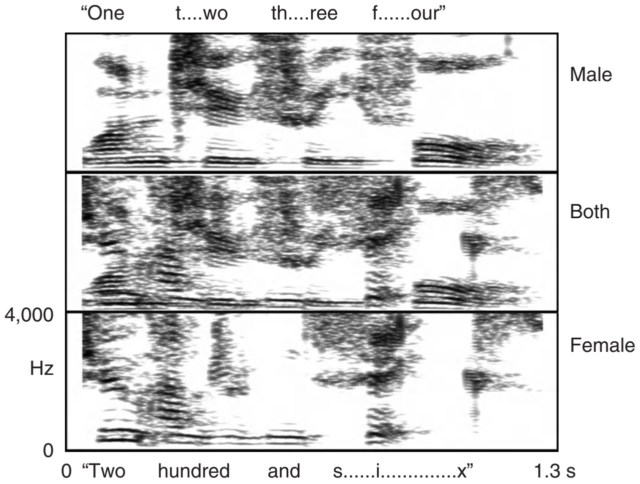


Figure 1 Narrow-Band Spectrograms

Notes: (Top) a male talker saying "One two three four"; (bottom) a female talker saying "Two hundred and six," and (middle) both utterances mixed together. The darker the spectrogram, the more energy there is in the sound at a particular frequency and time. The horizontal axis is time; the vertical axis frequency goes as high as 4,000 Hz. The roughly horizontal parallel thin dark lines denote the harmonics of the voice, closely spaced for the low-pitched male voice and more widely spaced for the higher-pitched female. The difference in pitch helps the brain perceive the two separate sounds in the mixture.

As an example of our ability to separate simultaneous sounds, Figure 1 shows spectrograms of three sounds. A spectrogram plots as a function of time how much energy is present in a sound at different frequencies; it broadly represents the pattern of activity that sound produces in the inner ear and that is signaled to the brain by the auditory nerve. The upper spectrogram is of a male voice saying "One two three four," the lower spectrogram is of a female voice saying "Two hundred and six," and the middle spectrogram is of those two sounds added together. When we listen to this mixture, we clearly hear the two voices as separate yet simultaneous sounds, and we have no difficulty in following a particular voice over time. Yet it is not obvious to the eye which parts of the mixture belong to one voice and which to the other. The brain uses a variety of cues to pull apart these two sounds from the mixture.

Simultaneous Grouping: Harmonic Structure and Onset Time

One of the cues that is useful for both speech and music is harmonic structure. Any sound apart from a pure tone (resembling a whistle) consists of many different frequencies. For a sound with a distinct pitch, such as a musical note, or a sung or spoken vowel sound, these frequencies will all be whole-number multiples (harmonics) of the fundamental frequency. So, for example, when the oboe plays the A above middle C to tune-up an orchestra, its sound contains frequencies that are all multiples of 440 Hz—it has a harmonic spectrum. If the oboe played the A an octave lower, then the fundamental would be 220 Hz and the harmonics would all be multiples of the lower fundamental of 220 Hz. The harmonic structure of the voiced parts of male and female speech is visible in the spectrogram of Figure 1. The harmonics appear as groups of roughly horizontal parallel lines. For the male voice, these are closely spaced, reflecting its low pitch; for the higher-pitched female voice, they are more widely spaced. The brain is adept at separating mixtures of differently spaced harmonics, as for example at the beginning of the mixed spectrogram where the low-pitched male voice and the high-pitched female voice are sounding simultaneously. If two voices are speaking or two instruments playing on the same pitch rather than on different pitches, it is much harder to separate them because their harmonic frequencies coincide.

Natural sounds that do not have a distinct pitch are less constrained in their frequency content: They may have a continuous range of frequencies rather than a set of discrete harmonically related frequencies. For example, the "f" of "four" in the male speech or the "s" of "six" are noisy, unvoiced sounds and so do not show harmonic structure—there are no parallel horizontal lines on the spectrogram. One of the cues that the brain uses to separate mixtures of sounds like these is the common onset time of a group of frequencies. The "t" of "two" in the male speech of Figure 1 is an example. The abrupt increase in energy of some of the frequencies in the sound mixture indicates that a new sound has started, and the brain then hears those frequencies that have abruptly increased in energy as a new sound in addition to the continuation of the sound that was previously present (the "un" of hundred" in the female voice). This principle has been called the *Old+New heuristic* by Albert Bregman. Harmonicity and common onset-time are examples of relatively simple, low-level properties of sound that the brain uses to help to segregate into separate sources the mixture of sounds that is occurring at a particular time.

Music exploits both our abilities and our inabilities to separate sound sources. Sometimes the composer wants us to hear separate sounds, at other times the composer wants us to hear different sounds blend into a single composite sound. Polyphonic music, such as a Bach fugue in which there are two or more separate melodic lines, strives to keep each of the different voices distinct to the listener; music students are taught the principles of "voice leading," which if followed help listeners keep the different voices of the music perceptually separate. David Huron has demonstrated the close links between these musical principles and the results of well-controlled psychological experiments on listeners' abilities to separate simple sounds. For example, voicing leading principles prohibit simultaneous pairs of notes from occurring in consecutive unisons, fifths, and octaves. Sounds at these consonant intervals tend to blend because they have many harmonics in common. For the unison, all the harmonics are in common; for the octave, every harmonic of the higher note coincides with every even harmonic of the lower note. The perfect fifth (for example, from C to G) has a frequency ratio of 3:2 so every third harmonic of the lower note is at the same frequency as every second harmonic of the higher note. When two sounds are played on less consonant intervals (such as the minor third), or if they start at different times, then they are much easier to keep perceptually distinct. Polyphonic music thus avoids sounds that have the same onset time and similar harmonic structures.

On the other hand, composers frequently capitalize on our inability to separate sound sources to create a novel timbre or sound quality that fuses two or more separate instruments. At the beginning of his *Unfinished Symphony,* Schubert merges the sound of the oboe and clarinet playing in unison to give a novel timbre that does not sound like either instrument. Played badly, out of time or tune, the timbres of the individual instruments emerge. Another example is the "12th" or "octave quint" organ stop, which produces a tone at an interval of a 12th (octave plus 5th, giving a frequency ratio of 3:1) above the note played on the keyboard. When coupled with a traditional stop, the "octave quint" stop thus adds a tone at triple the frequency of the original note. The harmonics of this higher note thus coincide with every third harmonic of the original and so simply modify the timbre without producing the impression of a separate pitch. Were the "12th" to be sufficiently out of tune (or out of synchrony) with the original played note, it would stand out as a separate note at the higher pitch and both notes would maintain their original timbre.

Common harmonic structure and a similar onset time are examples of heuristics that are used to group sounds into simultaneous wholes. They contribute to what Bregman in his influential 1993 book, *Auditory Scene Analysis,* describes as low-level grouping cues. These cues are simple properties that are more likely in components that originate from a common sound source than in those that have come from independent sources. Although their use does not guarantee a complete and appropriate organization of sounds, it provides a better starting point than the raw mixture.

An interesting current experimental question is to what extent the transformations that take place early in auditory processing (for example in the cochlea, the brain stem, and the mid-brain) facilitate such low-level grouping.

Sequential Streaming: Continuity

So far, this entry has looked at how the brain separates the mixture of sounds that is occurring at a particular time into appropriate groups that correspond to different sources. Another aspect of ASA is the problem of how we track a *particular* sound source across time to produce a single auditory stream. The problem is illustrated in an early paper from 1950 by George Miller and George Heise on the trill threshold. When two notes of similar timbre but different pitch alternate, the sequence can be heard in two ways: either as a trill—a single auditory stream rapidly varying in pitch, or as two separate streams each consisting of repetitions of one of the pitches. If the two notes are close in pitch, as is the case with a musical trill, a single, trilling, stream is heard. But if the notes are wide apart in pitch but still played rapidly, then the single stream splits into two, and you hear a sequence of repeating high notes, together with a sequence of repeating low notes. The faster the repetition rate of the notes, the closer together in frequency must the two pitches be to maintain the single stream, trill percept. The brain interprets large discontinuities in pitch in terms of different sound sources.

Abrupt discontinuities in other sound properties can also produce a similar split into different streams. David Wessel demonstrated such an effect with changes in timbre. The sounds he used are illustrated in Figure 2.

The upper panel A of Figure 2 illustrates the pitches of a sequence of notes played on the same instrument (solid lines). It is heard as a rapidly repeating ascending three-note figure. If the alternate notes are now played on a different instrument (dashed lines), the percept is dramatically changed. Now each instrument forms a separate stream. The rapid ascending figure is no longer heard, but instead each instrument plays a descending figure at half the original speed. A similar result can be obtained if the solid and dashed notes denote different spatial positions rather than different instrumental timbres.

In the 1970s, Bregman demonstrated an important consequence of this sequential streaming: Listeners find it much more difficult to tell the order of individual notes across streams compared to within streams. For example, in Figure 2, listeners would hear clearly whether the second note came before or after the third in the top panel, but would find it difficult in the lower panel because the two notes are heard in different streams.

Sequential streaming takes a few seconds to build up. Overcoming our initial default assumption of a single source requires accumulating evidence. A surprising result from Robert Carlyon and Rhodri Cusack's research is that attention is necessary for this accumulation to take place; if the sequence of tones is presented as an unattended background to a primary task, such as watching a video, no accumulation of evidence takes place. The result is surprising within Bregman's framework, which proposes that low-level grouping takes place pre-attentively, and has led to suggestions that unattended auditory material might undergo only limited separation into separate sources.

Schematic Knowledge

So far we have seen how relatively simple properties of sound can help to organize sound mixtures into their component sources or streams. But the brain likely uses more specific knowledge about sounds or sound types to help it to carry out this difficult task. A simple example involves listening to two simultaneous vowels spoken on the same pitch, starting and stopping at the same time, coming from the same direction in space. Here, there are no low-level grouping cues to help a listener to separate the two sounds, yet when asked to identify what the two vowels are, listeners' responses are well above chance. Knowing something about possible speech sounds helps the listener to assemble the appropriate components.

ASA and Hearing Impairment

Our ability to attend to different sounds that are present simultaneously generally passes unnoticed while we enjoy normal hearing, but those suffering from sensorineural hearing loss caused by a damaged or ageing cochlea are often aware of a disproportionate impairment in listening to one sound in the presence of others. Their difficulties arise partly from a loss of frequency selectivity in the cochlea, which increases the range of frequencies in the background sound that can mask the sound of interest. This loss of frequency selectivity, which is also present with cochlear implants, can in addition impair the brain's ability to use harmonic structure to separate different sound sources. Were hearing aids able to separate wanted from unwanted sounds, these problems would be reduced. Unfortunately, there is no effective general solution to sound separation either in a research implementation or in the much more demanding context of a wearable hearing aid. In practice, the problem is addressed by making modern aids directionally sensitive, so that sound coming from a particular direction (usually straight ahead) is amplified more than that from other directions. This approach can work well in a relatively uncluttered, dry acoustic environment, but is much less successful when even moderate amounts of reverberation are present because echoes of unwanted sounds will come from the amplified direction.

The difficulty of using spatial direction in generally useful ways in hearing aids is reflected in the difficulty that listeners with normal hearing show in exploiting some spatial differences to separate simultaneous sounds. For example, John Culling and Quentin Summerfield revealed a surprising inability of listeners to use the main localization cue in the horizontal plane (the difference in time of arrival of sounds at the two ears) to separate simultaneous sounds. If no other grouping cues are present, giving two sounds different interaural time differences does not help the listener to hear them as separate sounds or in different locations.

Computational Auditory Scene Analysis

Research during the last 20 years into ASA has proved challenging for theories addressing some of the classic issues in hearing such as pitch perception or sound localization. Theories that have been developed to cope with the perception of the pitch or direction of a sound when it is the only sound present often do not generalize well to the problem of determining the pitches and directions of multiple simultaneous sounds. The last 20 years have also seen the development of computational auditory scene analysis (CASA). A primary motivation behind CASA is the presently brittle nature of automatic speech recognition. Recognition algorithms generally use overall measures of the sound from the input microphone to match stored templates of words. Additional sounds corrupt these parameters and give poor recognition. Again, a general solution to the sound segregation problem would improve matters, but has proved elusive. However, the development of sophisticated CASA models and their integration with speech recognition algorithms is now leading to stimulating insights into how the low-level cues and schematic knowledge distinguished by Bregman might in practice be combined in the analysis of auditory scenes.

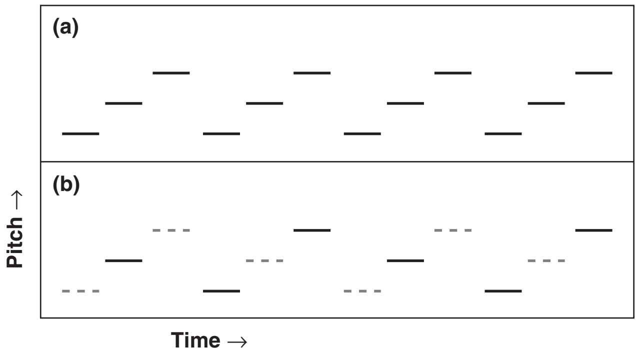


Figure 2 The Wessel Illusion

Notes: The upper panel (a) represents the pitches of a sequence of notes played on a single instrument. Listeners hear a repeating rapidly rising three-note figure. The lower panel (b) shows a similar sequence of pitches but with the even notes (dashed lines) played on a different instrument from the odd notes (solid lines). Listeners hear this sequence as two separate sound streams, each with a slowly descending three-note figure. The brain uses continuity of instrumental timbre to separate the mixture into two streams.

Chris Darwin

See also Audition: Pitch Perception; Auditory Localization: Psychophysics; Speech Perception

Further Readings

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge: MIT Press.

Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Science, 8*(10), 465–471.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25,* 975–979.

Cooke, M. P. (2006). A glimpsing model of speech perception in noise. *Journal of the Acoustical Society of America, 119,* 1562–1573.

Culling, J. F., & Summerfield, Q. (1995). Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay. *Journal of the Acoustical Society of America, 98,* 785–797.

Darwin, C. J. (1997). Auditory grouping. *Trends in Cognitive Science, 1*(9), 327–333.

Darwin, C. J., & Carlyon, R. P. (1995). Auditory grouping. In B. C. J. Moore (Ed.), *The handbook of perception and cognition: Vol. 6. Hearing* (2nd ed., pp. 387–424). London: Academic Press.

Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception, 19,* 1–64.

Miller, G. A., & Heise, G. A. (1950). The trill threshold. *Journal of the Acoustical Society of America, 22*(1), 637–638.

Wang, D. L., & Brown, G. J. (2006). *Computational auditory scene analysis.* Hoboken, NJ: Wiley.

Yost, W. A., Fay, R. R., & Popper, A. N. (Eds.). (2008). *Auditory perception of sound sources.* New York: Springer.

AUDITORY SYSTEM: DAMAGE DUE TO OVERSTIMULATION

Sound is the stimulus that activates the hair cells in the inner ear and starts the sequence of chemical and electrical processes that activate auditory nerve fibers and ultimately lead to perception. Sound that is too loud and lasts too long can severely damage the hair cells, can reduce or abolish activation of auditory nerve fibers, and can lead to impaired perception. This entry discusses damage to the auditory system caused by overstimulation.

Noise exposure levels are strictly regulated in the workplace to prevent hearing loss; typically a level of exposure of about 80 to 85 dB(A) is allowed for 8 hours, depending on the jurisdiction, and assumes the remaining 16 hours per day to be noise free. Every 3 dB(A) increase in sound level would halve the allowed exposure time. However, exposure to recreational noise in bars, sport venues, and concert halls may also lead to gradually accumulating hair cell loss and thus to hearing loss. For instance, the average noise levels in a hockey arena can be more than 100 dB(A) during the 3-hour game. Taking the 80–85 dB(A) industrial norm, this would then result in an allowable exposure time of only 5 to 15 minutes. Often, these games are watched in addition to the 8-hour allowable workplace noise, and the cumulative effect of these exposures can result in hair cell damage.

Hair cells come in two types, the inner and outer hair cells. Both inner and outer hair cells function as microphones, converting vibration into membrane potential changes. In addition, the outer hair cells function as mechanical amplifiers that increase both basilar membrane vibration amplitude as well as fluid motion near the inner hair cell stereocilia. Hair cell damage as a result of noise exposure typically starts with the outer hair cells in the high frequency region of the inner ear and, when the overstimulation is even more severe, is followed by loss of the inner hair cells in that same region. Complete loss of outer hair cells leads to the loss of amplification and produces about 40 dB hearing loss, and loss of frequency selectivity; combined, this results in deteriorated (speech) perception. When the inner hair cells are also damaged, the auditory nerve fibers that synapse with these cells commonly degenerate. In regions without hair cell damage, overstimulation also leads to excess inner hair cell release of the excitatory neurotransmitter glutamate, which in turn opens receptor channels in auditory nerve fiber terminals and allows excess calcium to flow into these terminals.

This causes damage to the inner hair cell auditory nerve fiber synapse resulting in loss of stimulation of the auditory nerve fibers and, hence, hearing loss. This neurotoxic process can be reversed under certain conditions, and the synapses may be fully repaired in less than a week. This neurotoxic damage is one of the substrates for the so-called temporary threshold shift following noise trauma. Hair cell damage, however, is irreversible and leads to a permanent threshold shift. Whereas the hearing loss can be compensated by amplification provided by hearing aids, the loss in frequency selectivity that results from outer hair cell loss cannot be ameliorated.

Noise exposure typically damages the outer hair cells in the high-frequency region of the cochlea, leaving the inner hair cells intact. Auditory nerve fibers that innervate the inner hair cells in regions with damaged outer hair cells typically have lower spontaneous- and driven-firing rates than do auditory nerve fibers innervating regions without any hair cell damage, which are typically tuned to low frequencies. Hence, the firing-rate profile across all auditory nerve fibers after noise exposure is no longer frequency independent as it is in normal ears, that is, is no longer balanced across the frequency range of hearing.

In several auditory cortical areas, there is a frequency-place map, also called tonotopic map. This map reflects the frequency analysis of the inner ear. An imbalance in the output of the auditory nerve fiber array results in changes in the levels of inhibitory and excitatory activity at the various stations in the central auditory system and ultimately leads to changes in the organization of the auditory cortex. Cortical neurons that would normally be most sensitive to the high frequencies that are affected by the hearing loss (e.g., above 6 kilohertz [kHz]) will after noise trauma become most sensitive to frequencies representing the edge of the hearing loss range (i.e., 6 kHz). As a result, a large part of the auditory cortex becomes sensitive to this edge frequency and this frequency becomes overrepresented in the tonotopic map. The same imbalance that gives rise to change in the frequency representation in the auditory cortex will also lead to increased spontaneous activity in the cortex and other parts of the central auditory system. This could be the cause for tinnitus (ringing in the ears), which often accompanies noise-induced hearing loss.

Damage to the auditory system is not restricted to hair cell loss. Degeneration of auditory nerve fibers underneath missing inner hair cells may subsequently occur; even as late as 8 months after the noise trauma. This degeneration is accompanied by newly formed synaptic endings in the cochlear nucleus. Therefore, noise-induced hearing loss may progress as a neurodegenerative disease with the capacity for synaptic reorganization within the cochlear nucleus. Any beneficial aspects of this reorganizational plasticity have yet to be shown.

Jos J. Eggermont

See also Audiology; Audition: Disorders; Auditory Frequency Analysis, Neural; Auditory Thresholds

Further Readings

Hodgetts, W. E., & Liu, R. (2006). Can hockey playoffs harm your hearing? *Canadian Medical Association Journal, 175,* 1541–1542.

Kim, J. J., Gross, J., Potashner, S. J., & Morest, D. K. (2004. Fine structure of long-term changes in the cochlear nucleus after acoustic overstimulation: Chronic degeneration and new growth of synaptic endings. *Journal of Neuroscience Research, 77,* 817–828.

Kujawa, S. G., & Liberman, M. C. (2006). Acceleration of age-related hearing loss by early noise exposure: Evidence of a misspent youth. *Journal of Neuroscience, 26,* 2115–2123.

Noreña, A. J., & Eggermont, J. J. (2005). Enriched acoustic environment after noise trauma reduces hearing loss and prevents cortical map reorganization. *Journal of Neuroscience, 25,* 699–705.

Noreña, A. J., & Eggermont, J. J. (2006). Enriched acoustic environment after noise trauma abolishes neural signs of tinnitus. *NeuroReport, 17,* 559–563.

AUDITORY SYSTEM: EVOLUTION OF

In the broadest sense, hearing is the detection, classification ("what"), and localization ("where") of mechanical disturbances of the surrounding medium. Such disturbances include sound—properly defined as longitudinal pressure waves propagating in an elastic medium—and other mechanical waves. Other sensory systems detect mechanical disturbances such as touch, vibration, and pressure sensors, and the origin of hearing is entwined in their evolution, but dedicated sound-receiving structures have only evolved in arthropods and vertebrates. The ancestral sensory cells in all hearing organs can be traced back to mechanoreceptors in organisms at least in the Cambrian (about 500 million years ago) period. This entry describes the evolution of the auditory system.

In insects, hearing organs developed independently at least 19 times from mechanoreceptors in the cuticle coupled to tympanic membranes. By contrast, the vertebrate inner ear originated only once. The vestibular structures that provide information about orientation and gravity are essentially unchanged during the course of evolution, but the hearing structures of the inner ear originally evolved for underwater hearing and show significant changes with the transition from water to land. In water, the similar density of water and tissue allows sound to travel through the body and stimulate auditory sense organs either directly or via mechanically coupled structures such as gas bladders. When vertebrates moved onto land, their ears may have resembled those of their fish ancestor's, responding to sound-induced vibrations of the skull (similar to bone conduction) and to substrate vibrations. On land, however, the impedance mismatch between the air and tissue causes most sound energy to be reflected, and physical constraints on detection and increased importance of detecting airborne sound probably led to the emergence of the eardrum (tympanum) and middle ear. Tympanic ears evolved during the Triassic period, probably when the spiracular openings were covered by a layer of tissue contacted by the middle-ear bone. Tympanic ears were previously assumed to have originated once, but the current consensus is that they emerged independently at least five times, in the lines leading to amphibians, turtles, lizards and snakes, archosaurs (crocodiles and birds), and mammals. Consequently, the auditory sense organs and nuclei in the central auditory pathways of land vertebrates are not necessarily homologous, but have emerged in parallel, subject to similar selective pressures for communication and localization.

In the early ear, the spiracles opened into the mouth cavity, leading to acoustically coupled eardrums. Thus, a sound wave that pushed in the eardrum on one side would travel though the mouth cavity and push out the eardrum on the other ear. These coupled ears can be extremely directional, as exemplified by modern lizard ears. The formation of a closed middle ear cavity in mammals and birds is a derived condition and would have profoundly changed the operation of the ear by decoupling the eardrums. Evolutionary processes that might have led to closed middle ears are numerous and may include requirements to shield the eardrums from respiratory noise. Additionally, although coupled ears provide exquisite directional information, they also exhibit nulls, or areas of low sensitivity. Whatever the cause(s) of the isolation of the middle ears, these changes would have led to a nondirectional ear, and selection for features that improve localization, such as external ears. External ears generate a robust, directionally dependent sound shadow, especially at high frequencies, and mammals have developed the three-ossicle middle ear with increased sensitivity to higher frequency sound, allowing improved detection of these sound shadows. On top of these peripheral changes, additional neural processing of interaural time and level differences would have been needed to compensate for the loss of directional information from the eardrum, leading to specialization for central processing of sound source localization information. Thus, many of the features of the avian and mammalian central auditory systems may reflect modifications of existing structures in already functioning land vertebrate systems. The increased sensitivity to airborne sound and higher frequencies may have provided the major evolutionary pressures for refinement of central auditory processing at a time when the major land vertebrate groups had already separated.

In the hindbrain, major differences among vertebrates are related to temporal and binaural processing. The auditory midbrain, the inferior colliculus or torus semicircularis, is a major processing site, especially in amphibians. Here, response properties such as selectivity for specific combinations of frequencies, specific temporal patterns, specific sound locations, or specific pulse-echo delays (bats) emerge. In birds and mammals, the greatest changes are in the forebrain auditory centers. Specialized communication systems based on vocal learning have evolved in parallel in humans, songbirds, parrots, and hummingbirds. Both humans and songbirds learn to produce their complex vocalizations early in life, exhibiting a strong dependence on hearing the adults they imitate, and have dedicated forebrain centers for vocal production and learning. Mammals are distinguished from other vertebrates by the evolution of the cortex in place of the more usual nuclear organization. The mammalian auditory cortex is characterized by both precisely tuned cortical systems with multiple auditory areas and specialized "where" and "what" pathways. Thus, the selection pressures that shaped the auditory system through evolution may be related to localization, communication, and specializations such as echolocation.

Catherine E. Carr and
Jakob Christensen-Dalsgaard

See also Animal Frequency and Pitch Perception; Audition: Temporal Factors; Auditory Processing: Central; Auditory Processing: Peripheral; Auditory Scene Analysis

Further Readings

Christensen-Dalsgaard, J. (2005). Directional hearing in non-mammalian vertebrates. In A. N. Popper & R. R. Fay (Eds.), *Sound source localization* (pp. 67–123). New York: Springer.

Clack, J. A. (1997). The evolution of tetrapod ears and the fossil record. *Brain, Behavior and Evolution, 50,* 198–212.

Grothe, B., Carr, C. E., Casseday, J. H., Fritzsch, B., & Köppl, C. (2005). The evolution of central pathways and their neural processing patterns. In G. A. Manley, A. N. Popper, & R. R. Fay (Eds.), *Evolution of the vertebrate auditory system* (pp. 289–359). New York: Springer.

Lombard, R. E., & Bolt, J. (1979). Evolution of the tetrapod ear: An analysis and reinterpretation. *Biological Journal of the Linnean Society, 11,* 19–76.

Manley, G. A., & Clack, J. (2004). An outline of the evolution of vertebrate hearing organs. In G. A. Manley, A. N. Popper, & R. R. Fay (Eds.), *Evolution of the vertebrate auditory system* (pp. 1–26). New York: Springer.

AUDITORY SYSTEM: STRUCTURE

When the ancestors of land animals first left the sea, they faced many new challenges; one of which was the detection of airborne sounds. Animals have many types of vibration receptor but, like most cells in the body, they require a fluid environment. This posed a problem because, when sound waves meet a solid or liquid, most of the sound energy is reflected and less than 10% is transferred into the solid or liquid. This problem (impedance matching) has been overcome by the evolution of structures that take the highly dispersed energy of the air molecules and focus it onto a small solid point to produce vibrations in the fluid around the specialized vibration receptors. This entry describes the structure of the ear and the pathways within the central nervous system that process its input.

Structure

The external ear, or pinna, acts as a funnel directing sounds into the external auditory canal to produce movement of the eardrum (tympanic membrane). The pinna can become greatly enlarged as in the long-eared bat, and in species such as the horse or cat, the pinna is independently movable and can be directed toward the source of a sound. Human pinna movement is vestigial but the pinna still generates a characteristic notch in the intensity of higher frequencies. The distinctive ridges and curves of the pinna form reflective surfaces that lead to some frequencies having a longer pathway into the ear canal. This introduction of a small delay or phase change can produce a waveform cancellation and a reduction in sound level among some high frequencies. The precise frequencies that are reduced in the notch are altered by the direction of the sound source, and this is thought to be important in estimating sound elevation and in sound localization for people with unilateral deafness.

The tympanic membrane is stretched taut across the auditory canal (see color insert, Figure 4). Pressed against it is the first of three tiny bones (the ossicles) that act as a lever to amplify and focus the force exerted by the air molecules. These bones are the malleus (mallet), incus (anvil), and stapes (stirrup). The malleus is attached to the eardrum while the stapes is pressed against the oval window: a membrane-covered opening in the fluid-filled inner ear. The long arm of the malleus takes the weak movements of the tympanic membrane and transfers them to the small base of the stapes in a process that amplifies the pressure by almost 15 times. The joints between the ossicles are conventional ball and socket joints and are susceptible to pathological processes and the base of the stapes can become fused to the wall of the inner ear (otosclerosis) and produce a conductive hearing loss. In conductive hearing loss, the ear is still sensitive to vibrations that come directly through the bone, and so a person would still be able to hear a train from several miles away if the person pressed his or her head against a railway track. This type of hearing impairment is more easily treated than is sensorineural hearing loss, which involves the permanent loss of the hair cells of the inner ear that detect the vibrations, or the nerve cells that contact them. The actions of the ossicles are controlled by tiny muscles called the tensor tympani (to malleus) and the stapedius (to stapes). These muscles can act via the stapedius reflex to reduce the amplitude of the vibrations produced by loud, low-frequency sounds. Paralysis of the stapedius muscle can produce hypersensitivity to sounds.

The sensory structures of the inner ear are contained in a labyrinth of interconnected tubes within the dense bone at the base of the skull. The sensory cells for hearing are located at the base of one of the membrane-bound tubes and form the organ of Corti. This spirals around for two and a half turns in the part of the labyrinth called the cochlea. The mechanical properties of the membrane on the floor of the organ of Corti (the basilar membrane) mean that the membrane at the base of the cochlea is taut and narrow (blue line) and tuned to high frequencies, and the wider, more floppy membrane at the apex (red line) is tuned to low frequencies (see color insert, Figure 5). This is counterintuitive because the base of the cochlea, which resembles the shape of a snail shell, would be expected to have the wider part of the membrane. The membrane at the base is actually stretched across a narrower gap because there is a spiral shelf of bone protruding from the center of the cochlea. This shelf is broad at the base and narrow at the apex.

Sitting on the basilar membrane are a single, spiral row of inner hair cells (IHC) and three rows of outer hair cells (OHC). The IHC are the sensory receptors that transform (transduce) the mechanical sound energy into neural impulses via synapses with type I nerve fibers that make up about 90% of the cochlear sensory fibers. Progressive changes in the width and tension of the basilar membrane mean that the IHC at any particular point are most sensitive to one particular frequency. This means that the frequency components making up the sound wave are coded by the position of the nerve endings contacting IHC as well as by the rate at which the hair cell vibrates. The OHC appear to function primarily as a cochlear amplifier to modify the properties of the basilar membrane and actively increase the amplitude of its vibrations by changing their length in time with the vibrations of the membrane. Thus, the cochlea is not a passive listening device but has a feedback mechanism for actively amplifying small vibrations. The OHC are contacted by some thin, type II sensory nerves but these are much less numerous than the type I nerves.

Pathways

The sensory nerve fibers have their cell bodies in channels in the conical bone at the center of the cochlea (the modiolus) and form the spiral ganglion. The central processes of these cells form the cochlear branch of the VIIIth cranial nerve and enter the brain where they immediately split into ascending and descending branches. The ascending branch of the type I fibers forms specialized synapses on the spherical bushy cells (SBC). These are giant synapses that engulf the spherical cell bodies and permit secure and accurately timed synaptic transmission. The descending branches also produce giant synapses around the cell body of globular bushy cells (GBC) and regular synapses with the dendrites of other, usually larger, nerve cells.

The primary nerve fibers enter the cochlear nucleus in an ordered sequence with the low-frequency (red, LF) fibers being more ventral and lateral and the high-frequency (blue, HF) fibers being more dorsal and medial. This regular frequency gradient is maintained at all levels of the ascending auditory pathway right up to the cortex.

Small cells such as the bushy cells, with their small "bush" of stunted dendrites, are all sharply tuned as they are primarily driven by a single sensory fiber. By contrast, some large nerve cells have widely branching dendrites and receive inputs from many fibers that represent a wide range of frequencies. This means that they can simultaneously compare the inputs over a wide spectral range and possibly code for changes in the relative intensity of different frequency components within a sound (important for determining the elevation of a sound). By synapsing on a variety of output neurons in the cochlear nucleus, the input from the sensory nerve is split up into at least five different output pathways. As cells in the cochlear nucleus also receive inputs from the auditory cortex, the contralateral cochlear nucleus, and vestibular and somatosensory nuclei, along with intrinsic excitatory and inhibitory neurons, their output has already undergone significant central processing.

Most of the output from the cochlear nucleus is destined for the main auditory nucleus of the midbrain (the inferior colliculus), but the routes by which the information get there are varied and reflect the importance of binaural (two-ear) comparisons. Some cells project directly to the inferior colliculus, but most are involved in indirect pathways. Two of these involve the small bushy cells whose outputs are involved in localizing low-frequency sounds. Sounds off to one side of the head will arrive at the near ear slightly sooner than at the far ear, and by comparing the time of arrival at the two ears, the brain can determine if the sound is coming from directly ahead or is off to one side. The two types of bushy cells both project to the brainstem structure called the superior olivary complex (SOC), which in turn projects to the inferior colliculus. The spherical bushy cells project bilaterally to the medial superior olivary nucleus (MSO) and the recipient cells act as coincidence detectors to compare the timing of the signals from the two ears. These cells are sensitive to differences as short as 20 millionths of a second even though the action potentials that carry the neural information are each about one thousandth of a second long. This is rather like using a grandfather clock with a one-second pendulum to measure the time taken by athletes in a time trial! By marking the position of the pendulum, you could measure times much shorter than a second, and in the same sort of way, the MSO cells will only fire an action potential if the inputs from the two cochlear nuclei are coincident. If the action potentials from the two input cells arrive at slightly different times, then they can cancel each other out and the MSO cell will not fire. The globular bushy cells also project bilaterally to the superior olivary complex: to the lateral nucleus (LSO) on the same side but to an adjacent nucleus (the medial nucleus of the trapezoid body or MNTB), that contains inhibitory cells, on the opposite side. These inhibitory cells project to the adjacent lateral nucleus. Thus, the cells of the lateral nucleus also act as binaural comparators to compare the arrival time of input from the two ears. In this case, coincident arrival of an excitatory and an inhibitory input means that the two inputs cancel each other out. If the inhibition is delayed slightly, then the LSO cell has time to fire before the inhibition takes effect.

The information contained in the different ascending pathways is selectively recombined within the central nucleus of the inferior colliculus. This nucleus is organized into laminae that are stretched out at right angles to the frequency gradient. Thus, the top lamina in the central nucleus responds to low frequencies, and the deeper laminae respond to successively higher frequencies. Within the laminae, there are both excitatory and inhibitory cells, and most project to the thalamus.

The main auditory nucleus in the thalamus is called the medial geniculate body and it has three main divisions: medial (MGM), dorsal (MGD), and ventral (MGV). The ventral nucleus receives the bulk of the output from the central inferior colliculus and projects to the core auditory areas that include the primary auditory cortex and adjacent areas that maintain a representation of the frequency gradient in the cochlea. These core areas then project to the surrounding belt areas, and in primates at least, the belt areas then project to parabelt areas. Different cortical areas have separate functions and the specialized human language areas may correspond to the parabelt. In the rhesus monkey, three core areas receive parallel input from the MGV, and there are parallel processing streams between cortical areas. Thus, the belt and parabelt areas project to multimodal areas of the frontal and temporal lobes in separate streams.

The ventral stream ("what" pathway) is more concerned with sound identification, and the dorsal stream is more concerned with sound localization ("where" pathway).

All neocortical areas have six layers, and the only types of projection neurons are the excitatory pyramidal cells (PC). Cells located in the deepest layers (V and VI) project back down to the auditory thalamus and these descending fibers outnumber the ascending ones by about four to one. The layer V pyramidal cells also project to the dorsal and lateral nuclei that form the periphery of the inferior colliculus as well as most other auditory structures including the SOC and the cochlear nucleus. The projections to the SOC appear to contact scattered cells that project back to the cochlea and modulate the activity of the inner and outer hair cells. Thus, powerful descending pathways arise in the cortex as well as in the thalamus and inferior colliculus that provide feedback and can potentially affect hearing sensitivity and mediate the effects of selective attention.

Mark N. Wallace

See also Auditory Localization: Physiology; Auditory Processing: Central; Auditory Processing: Peripheral; Auditory Receptors and Transduction

Further Readings

Kaas, J. H., & Hackett, T. A. (2008). The functional neuroanatomy of the auditory cortex. In A. I. Basbaum (Ed.), *The senses: A comprehensive reference* (Vol. 3, Chapter 45, pp. 765–780). Boston: Elsevier. DOI: 10.1016/B978–012370880–9.00057–8

Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences, 26,* 100–107. DOI: 10.1016/S0166–2236(02)00037–1

AUDITORY THRESHOLDS

The term *auditory threshold* usually means the minimum sound pressure level needed to sense the presence of an acoustic stimulus. Most of us are familiar with this type of auditory threshold through the experience of having a hearing test. In the most common type of hearing test, the minimum sound pressure level (SPL, specified in decibels or dB) needed to detect the presence of pure tones (sine waves having clear pitches) is measured for tones varying from low to high frequencies (frequency is the physical attribute of sound most closely corresponding to perceived pitch). Usually, this type of hearing test is performed while the listener wears earphones so that each ear can be tested separately.

Representative population values for the minimum detectable SPL for pure tones presented through earphones have been obtained by testing large groups of listeners thought to have normal hearing sensitivity. These normative values (called minimum audible pressures) reveal that humans are most sensitive to sounds in the middle range of hearing (e.g., from about 1–4 kilohertz [kHz]) and less sensitive to low- and high-frequency sounds. The general shape of this function is largely determined by the acoustics of the outer ear and the power transfer characteristics of the middle ear, although other factors such as biological noise (heart beat, respiration, etc.) can influence those values. The plot of normal thresholds in SPL as a function of frequency resembles a somewhat asymmetric bowl shape with a low-frequency slope that decreases fairly gradually (roughly–20 dB per octave or doubling of frequency), whereas the high-frequency slope is much steeper (as much as 80 dB/octave). Furthermore, the location in frequency of the high-frequency slope is highly dependent on age and may decrease substantially even in the early years of adulthood. This apparently "normal" decrease in high-frequency hearing sensitivity has been the basis for several recent commercial products (e.g., high-frequency ring tones) that emit high-frequency sounds audible to teenagers and young adults but not older adults.

The audiogram that is generated by the type of hearing test described is taken as an estimate of a person's auditory sensitivity. The normative values of SPL are the basis for hearing threshold levels (HTLs) such that the just-detectable SPL across frequency is set to 0 dB HTL essentially flattening out this bowl-shaped function. Thus, on the audiogram, thresholds are plotted in reference to these normative values. Hearing loss results in auditory thresholds that are elevated relative to normal (although paradoxically an audiogram plots increases in sound level as decreases along the ordinate). The most common types of hearing loss result from infections of the middle ear or in pathology affecting the sensory cells of the inner ear. Comparing thresholds for sounds presented by earphones with those presented by a bone-conduction vibrator often provides an indication of which type of hearing loss is present. This entry describes absolute, differential, and other types of thresholds and how to make sense of threshold levels.

Absolute Threshold

Many other types of thresholds for acoustic stimuli are referred to in the auditory literature. One helpful distinction in sorting through these various threshold measures is that of an absolute threshold versus a relative or differential threshold. The type of threshold mentioned earlier that would typically be measured by an audiologist administering a hearing test is an example of an absolute threshold. No sounds other than the pure-tone signal are presented to the test ear, and the judgment made by the listener is simply whether the tone is present or absent. Before discussing differential thresholds, it is useful to describe what the term *threshold* itself actually means.

A threshold is usually expressed as a single number. So, continuing the previous example, if an audiologist measures a person's threshold for a 500-hertz (Hz) pure tone, a single number is usually recorded; for example, 20 dB SPL (SPL will continue to be used in the example while acknowledging that an audiogram is plotted in HTL). It is almost never the case, however, that the 500-Hz tone would always be heard if it were played at 20 dB SPL or that it would never be detected if it were presented at 19 dB SPL. Detection thresholds usually refer to the point where a stimulus is detected an intermediate percentage of the time. More commonly, the percentage of times a listener detects the presence of an auditory signal varies with SPL.

Suppose one measured the percentage of times a listener detected a tone at five different SPLs, let's say 16, 18, 20, 22, and 24 dB SPL. The tone normally would be presented several times at each level so that a reliable estimate of performance is obtained. At 16 dB SPL, the listener detected the tone 10% of the time; at 18 dB SPL, it was detected 25% of the time; at 20 dB SPL, 50% of the time; at 22 dB SPL, 75% of the time; and at 24 dB SPL, 90% of the time. The mathematical expression describing how performance changes as SPL is varied is called a psychometric function. So, rather than there being a single value of SPL dividing performance between inaudible and audible sounds, the likelihood that the signal was detected is different for different presentation levels. Thus, the *probability of detecting* the presence of the signal is a function of signal SPL. Any of the levels mentioned could be considered a "threshold" because the signal was detected at that level at least some of the time. However, there is usually an agreed-upon criterion value—such as the example of the SPL producing 50% correct used previously—that yields the single number reported for "threshold."

In the earlier example, little information was given about the way the test was administered. However, it is easy to imagine that different listeners might react differently depending on their own tendency to respond one way or the other when they are uncertain about their answers. Suppose the listener was told that the signal would be presented during the interval of time a light was illuminated and that many such "trials" would be tested. Furthermore, the signal tone would only be present on one-half of the trials and the probability it would occur on any given trial was 0.5. After each trial, the listener must answer "yes" a tone had been presented or "no" it had not. It is easy to imagine two listeners whose sensitivity to the test tone is the same but their response behavior is different. One listener, for example, feels that it is important never to miss a signal when it is presented. This listener might be inclined to respond "yes" even if not sure that a tone had been presented. This type of behavior or *response bias* would lead to detecting the signal most of the time it was presented (a high proportion of "hits"), but it also would lead to a relatively large number of "yes" responses when no signal was presented. Those incorrect "yes" responses are called *false alarms*. The second hypothetical listener has a much more conservative approach and only votes "yes" when he or she is absolutely sure the signal was present. This listener would probably have few false alarms but would also miss the signal when it was present on more presentations than the other listener. Thus, specification of performance beyond simple percent correct should consider listener response behavior such as computing the ratio of hits to false alarms. A common measure or index of the detectability of a signal is referred to in *signal detection theory* as d' (pronounced "d prime").

Differential Threshold

Now we return to the idea of a relative or differential threshold. Suppose we were interested in the smallest difference in frequency a listener could discriminate. Let's assume that the stimuli are pure tones and that one tone is 2,000 Hz and the other tone is 2,000 Hz plus some small increment in frequency. The listener hears the two tones in sequence one in each of two observation intervals of a trial with the order of presentation chosen randomly. The listener then must indicate which of the two observation intervals contained the tone that was higher in pitch. If the difference in frequency between the two tones is small, the listener may not be able to tell which one is higher and simply must guess. In that case, the percent correct performance may be no better than chance (i.e., 50% correct). However, if the difference in frequency between the two tones is increased, performance should improve with percent correct scores increasing above chance until, with a sufficient difference in frequency, performance may approach 100% correct.

As with the previous example, we can specify a percent correct point on this psychometric function to define as the criterion performance for "threshold." For example, it could be that a frequency of 2,020 Hz was discriminable from 2,000 Hz 75% of the time. That value is considered a relative or differential threshold. It is also called a just-noticeable difference (JND) or difference limen (DL). Another common expression for a differential threshold is the ratio formed by the JND relative to the standard stimulus, which in this case would yield 20 Hz/2,000 Hz or 0.01. This ratio is also called a Weber fraction after the famous psychophysical scientist who proposed that this ratio was a constant value for many types of sensory discriminations. Many types of differential thresholds are reported in the auditory literature reflecting the discriminability of various aspects of simple tones or of more complex stimuli such as modulated tones or speech.

Other Types of Thresholds

As a final illustration of the various types of auditory thresholds, consider the following. A listener is presented with a sequence of individual words that initially are so soft as to be inaudible. The task of the listener is to report whenever a word is heard. The listener does not have to understand the words, only indicate that he or she was aware a sound had been presented. As the level of the sounds is raised, the probability of detection increases. As with the two earlier examples, detection performance would take the form of a psychometric function that varied from chance to nearly perfect performance. And, it would be possible to specify an SPL yielding a specified level of performance that we could call a speech detection threshold.

Now suppose the listener is not asked to indicate detection but instead must identify the spoken words. Typically, the level of the sound would have to be raised a few decibels above the just-measured threshold for speech detection to find the point where the speech was understandable a specified percentage of the time. So, a second psychometric function is obtained for the same stimuli, but this time for the task of speech recognition. Here we have two psychometric functions for the same sounds but for different tasks: one for detection and the other for recognition. Finally, suppose this speech detection/recognition experiment is repeated but now the words are not presented in quiet but instead a continuous "white" or Gaussian noise is played at a moderate SPL. Now the speech would not be detectable or recognizable at the same SPLs as before but instead the level of the speech would have to be increased so that the words could be heard above the noise. The psychometric functions would be shifted to these new levels, and the threshold points obtained from the psychometric functions would be considered "masked thresholds" because the noise acted to mask the speech signals. Thus, the values of threshold obtained for a given stimulus—speech in this example—may be affected by the presence of other sounds.

Making Sense of Threshold Values

To summarize, although the most familiar auditory threshold is for detection of pure tones, many other types of thresholds for different stimuli and tasks are possible and occur frequently in the auditory literature. Many variables affect the threshold values obtained in an experiment or clinical procedure, including the nature of the stimulus, the experimental technique, the instructions given to the observer, and the specific task. To make sense of the threshold values (particularly for comparisons across laboratories and clinics) that are obtained requires careful specification of these variables and how they were controlled.

Gerald Kidd, Jr.

See also Audiology; Audition: Disorders; Auditory Processing: Peripheral; Signal Detection Theory and Procedures

Further Readings

Green, D. M., Kidd, G., Jr., & Stevens, K. N. (1987). High-frequency audiometric assessment of a young adult population. *Journal of the Acoustical Society of America, 81,* 485–494.

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics.* Huntington, NY: Robert E. Krieger.

Killion, M. C. (1978). Revised estimate of minimum audible pressure: Where is the "missing 6 dB"? *Journal of the Acoustical Society of America, 63,* 1501–1508.

Rosowski, J. J. (1991). The effects of external- and middle-ear filtering on auditory thresholds and noise-induced hearing loss. *Journal of the Acoustical Society of America, 90,* 124–135.

Yost, W. A. (2007). *Fundamentals of hearing: An introduction* (5th ed.). San Diego, CA: Academic Press.

B

BAYESIAN APPROACH

The *Bayesian approach* to perception focuses on whether the computations that the brain performs in perceiving the world can be described as forms of Bayesian inference. *Bayesian inference* is the process of optimally extracting information from noisy inputs and, if needed, combining this information with prior knowledge. The Bayesian approach to perception typically involves mathematically precise modeling of perceptual experiments in humans. Central to Bayesian inference is *Bayes' rule* (also called *Bayes' theorem*), named after the 18th-century British mathematician and Presbyterian minister Thomas Bayes. Bayes' rule is a general equation directly derived from the basic tenets of probability theory, but in the context of perception, it is based on the fact that the brain's observations of physical stimuli (such as the orientation of a line or the identity of an object) are noisy. The rule states that the probability that a stimulus had a particular value given a set of noisy observations (the *posterior probability* of the stimulus) is proportional to the probability that a stimulus of that value generated those observations (the *likelihood* of the stimulus), multiplied by the probability of the stimulus value independent of the observations (the *prior probability* of the stimulus). In mathematical notation,

$$p(\text{stimulus} \mid \text{observations}) \ \alpha \ p(\text{observations} \mid \text{stimulus}) \times p(\text{stimulus}).$$

In this equation, p stands for probability, "|" means "given," and "α" means "proportional to." (In more abstract notation, Bayes' rule is written as $p(\mathrm{A}|\mathrm{B}) \ \alpha \ p(\mathrm{B}|\mathrm{A}) \ p(\mathrm{A})$, where A and B are random variables that could refer to the stimulus and the observations, respectively.) This entry discusses several concrete examples. Ernst Mach and Hermann von Helmholtz were among the first to apply the idea of Bayesian inference to sensory perception, although a similar notion can be found in the writings of Pierre-Simon Laplace. Strong experimental evidence for Bayesian inference in perception started emerging only in the 1980s, and some of it is discussed here. The Bayesian approach to perception should not be confused with Bayesian data analysis, which is an area of statistics with applications in neuroscience.

Perception as Information Extraction

Our percept of the world is a complex composite of pieces of uncertain knowledge extracted from noisy and ambiguous sensory inputs. The brain is different from a recording device such as a camera, which accurately captures the light in a visual scene but will not be able to tell, for instance, whether there is an animal in the scene. Indeed, in computer vision, sophisticated algorithms are used to extract that information. To humans and other animals, such higher-level knowledge can be of great ecological importance—for example, an animal in the scene could be a predator. To extract information from sensory inputs, the brain has to perform many different manipulations, including combining signals that come from the same source, making decisions, planning movements, allocating attention, and storing and retrieving information. Manipulations of sensory inputs that aim to extract behaviorally relevant information can be called *perceptual computations.* A key question in neuroscience and cognitive science is what algorithms and mechanisms the brain has developed to perform perceptual computations in the presence of uncertainty.

Uncertainty in Perception

Sensory information is uncertain. A lion chasing a giraffe must judge its position, speed, and direction of motion to decide when and how to attack, but because the giraffe is running fast, these judgments come with large uncertainty. A basketball player trying to pass the ball must keep track of the positions of his teammates and of the members of the opposing team, sometimes only seeing them out of the corner of his eye. When someone approaching you looks familiar, you need to use visual information to decide whether to greet your friend or avoid an awkward moment. When you are hiking in the forest and need to decide whether to jump over a stream, you have to judge the width of the stream as well as how far you can jump. When you are driving on the freeway and the car ahead of you brakes, you determine how to react based on uncertain variables such as your car's distance to it. The central idea of the Bayesian approach is to acknowledge this omnipresent uncertainty and ask how the brain can consider it optimally when performing perceptual computations.

An Analogy: The Apple Tree

Because sensory information is noisy, the brain constantly has to make guesses about the state of the world. Some guesses are better than others. Computing the goodness of possible guesses is known as Bayesian inference. Imagine you are shown the locations of some apples lying scattered in a field. You are told that the apples fell from a tree recently and are asked to guess the location of the tree. Your best guess will probably be somewhere in the middle between the apples because you know that apples don't fall far from the tree and tend to fall roughly evenly in all directions. You will be able to report a best guess of the tree location, and you will have an idea how probable each other location in the field is. For example, the farther outward a location is, the less you will believe that the apple tree was there. This changes if you know that there was a strong wind blowing from the west when the apples fell. Your best guess about the location of the tree would then be more to the east than it was before, and probabilities assigned to other locations will change accordingly. If you have additional information, for example that some areas of the field are swampy, you will be able to adjust your probabilities, because you know apple trees tend to grow on solid ground.

This example illustrates that one can use noisy observations (the locations of the apples) to obtain the probability of each of their possible causes (apple tree positions). However, for this to work, you need to know how a given cause generates observations (for example, that apples don't fall far from the tree). You might also have prior information about causes (such as that apple trees tend to grow on solid ground). Using these pieces of knowledge when you are given a particular set of observations will allow you to infer the goodness of possible guesses about the stimulus. This is exactly the logic expressed by Bayes' rule—in this example, it would state that the probability of an apple tree location given locations of apples in the field can be computed by multiplying the probability of those apple locations given that tree location with the prior probability of that tree location. This multiplication is followed by a *normalization,* which is the division by a constant factor to ensure that the total probability of all possible causes is equal to 1—a defining property of a probability distribution. One can write this as $p(\text{tree location} \mid \text{apple locations}) \propto p(\text{apple locations} \mid \text{tree location}) \times p(\text{tree location})$.

Bayes' rule is not just about apple trees, however. It is general and widely used in fields as different as statistics, economics, computer science, computational biology, and neuroscience. Moreover, it has practical applications, such as in Bayesian spam filters, which calculate the probability that an e-mail is spam based on the words it contains. These filters multiply the probability of finding those words in a spam e-mail by the prior probability that any e-mail is spam, followed by a normalization: $p(\text{spam} \mid \text{words}) \propto p(\text{words} \mid \text{spam}) \times p(\text{spam})$.

Bayesian Inference in Perception

This entry now discusses what Bayes' rule means in the context of perception. The apple tree in the previous example is analogous to a feature of an event or object in the world, often called the *stimulus* (in a somewhat confusing terminology, the event or object itself is sometimes also called the stimulus. This can be a simple variable such as the orientation of a line, the color of a surface, or the direction of motion of an object. It can also be more complicated, such as whether a desired object is present in a scene, what your coworker just said, the identity of the person approaching you, or how safe it is to change lanes on the freeway.

The apple locations are analogous to observations of the stimulus made by the brain. There are different ways of conceptualizing these. In their simplest form, the observations are variable measurements of the stimulus. For example, when the stimulus is a line oriented at 60°, a measurement of its orientation on a single trial could be 58° or 61°. This is sometimes called the *internal representation* of the stimulus. It is often assumed that internal representation of a stimulus is distributed around the true value of the stimulus, according to a roughly Gaussian (normal) distribution. Alternatively, the observations could be the variable activities of neurons in a population responding to the stimulus. For example, if the stimulus is an oriented line, then the observations could be the numbers of spikes emitted by the neurons in the primary visual cortex (V1) that are sensitive to orientation at that location. These numbers vary because you do not get the same pattern of response every time you present the same stimulus. Which description of the observations is most appropriate depends mostly on whether you study a perceptual phenomenon from a behavioral or a neural point of view: Psychologists tend to use noisy measurements, whereas neuroscientists tend to use neural activities. Most Bayesian models of perception are at the behavioral level.

Now that we have described what the stimulus and the observations are, we can formulate the central hypothesis of the Bayesian approach in perception: In many perceptual tasks, the brain computes the posterior probability distribution over the relevant stimulus variable based on the observations it has available, as expressed in the equation on page 201. One can think of this equation as a formalized description of a key task that the brain faces: to infer aspects of the world based on noisy sensory input. The lion infers in which direction the giraffe is moving, the basketball player where his teammates are, and the forest hiker whether he can jump over the stream. From the point of view of the observations, $p(\text{observations} \mid \text{stimulus})$ is sometimes called the *noise model,* the *variability* in the observations, or their *generative model.* In the apple tree analogy, the clue that there is a strong wind from the west is an element of the generative model. In the example of the oriented line, the Gaussian distribution or the form of the variability of the population of V1 neurons constitutes the generative model. When the observations are the activities of a population of neurons, one can think of that population as *encoding* the probability distribution over the stimulus (see Figure 1).

The posterior probability distribution, $p(\text{stimulus} \mid \text{observations})$, reflects the uncertainty about the stimulus: If this distribution has a single sharp peak, it means that based on the observations, only one of a few stimuli are plausible, and uncertainty is small. For example, when you are driving on the freeway on a sunny day and there is another car a short distance in front of you, the posterior probability distribution over your distance to it will be narrow. If, on the contrary, the distribution has a flat and broad peak, or multiple peaks of similar height, it means that based on the observations, many stimuli are plausible, and uncertainty is high. For example, if the same car is far ahead of you at nighttime, the posterior probability distribution will be broad. This is an important point because it means that, through Bayes' rule, the brain can encode not only the most likely value of a stimulus (compare: the most likely location of the apple tree, the most likely distance to the other car), but also the uncertainty about that value (compare: how large the area is where you believe the apple tree could still stand, how wide a range of distances to the other car is plausible).

Bayesian inference in perception is seldom a matter of perceiving a single stimulus. As already discussed, perceptual computation can involve elaborate manipulation of sensory information,

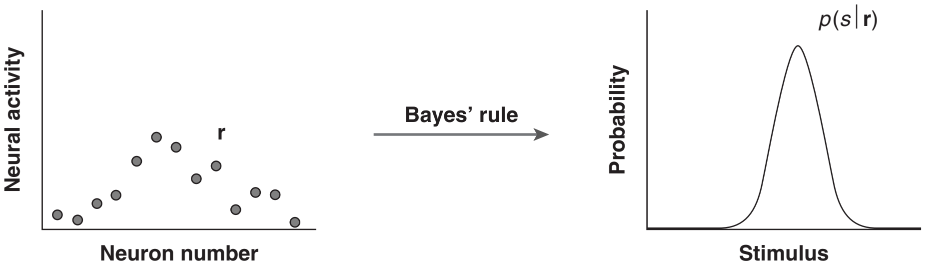


Figure 1 Schematic Illustration of Bayes' Rule in Perception (Neural Version)

Notes: A stimulus (e.g., orientation of a line) elicits activity in a population of neurons (left), whose activities (e.g., spike counts) we collectively denote by **r**. Based on these noisy observations, Bayes' rule computes the probability distribution $p(s|\mathbf{r})$ over the stimulus (right), providing the most likely value of the stimulus, as well as its uncertainty. One can think of the probability distribution being encoded in the pattern of activity.

which in many cases has to be done in multiple stages. For example, when you are deciding whether there is a predator hiding in the bushes, your brain has to infer the features of objects in the scene, as well as make the higher-level judgment about animal presence. In such situations, Bayesian inference can become complex, but the key feature remains that at every stage, a probability distribution is computed over the variables at that stage. Using Bayes' rule to compute with uncertainty and probability during a perceptual task guarantees the best possible performance; therefore, Bayesian inference is often referred to as *optimal inference,* or "optimally taking into account uncertainty." The Bayes-optimal strategy ensures that an organism makes fewer mistakes and is more efficient in perceiving the world and generating behavior.

Experimental Evidence

Perception researchers have found a large amount of evidence that the brain performs Bayesian inference, that is, computes posterior probability distributions over stimuli. This evidence is commonly provided by showing that those probabilities are used by humans in solving certain behavioral tasks. No experimental paradigm has contributed more to this than that of cue combination. The basic idea of this perceptual computation is that an observer is presented with two or more cues providing information about the same stimulus, and has to estimate the stimulus. For example, you might have a cat that likes to hide in the backyard. To find her, you use the movement you see as well as the faint meowing you hear. Because of noise, both in the world and inside the brain, cues from the same source are often not in complete agreement. For example, the best guess about the location of the cat based on its meowing may be slightly different from the best guess based on its image. One way to combine such cues would be to take the average of these guesses. However, this is not a good idea when one cue is much less certain than the other: if the cat is barely visible, the best combined guess may be close to the one based on what you hear. Indeed, the formalism of Bayesian inference applied to cue combination predicts that the best combined estimate is a specific combination of the individual guesses that gives less weight to the less certain cue. This prediction has been tested in a wide variety of cue combination experiments in which the disagreement between two cues is artificially varied. The general conclusion has been that humans do consider uncertainty in a close to Bayes-optimal way when combining cues. Evidence for Bayesian inference has been found in many other areas of perception as well, including decision making, visual motion estimation, color perception, and sensorimotor behavior.

Neural Basis

Even though Bayesian inference has become a fixture in behavioral descriptions of perception, the neural basis of Bayes-optimal computations has only recently started to receive scrutiny. The challenge here is to discover which operations the brain performs on neural activity so as to produce behavior that is close to Bayes-optimal. This problem is directly related to one of the central questions in neuroscience, namely that of the relation between brain and behavior. Eventually, predictions from this approach might be tested in physiology. Research into Bayesian inference in perception is expected to be of wide interest in the coming years.

Wei Ji Ma

See also Computational Approaches; Multimodal Interactions: Neural Basis; Multimodal Interactions: Visual–Auditory; Neural Representation/Coding; Signal Detection Theory and Procedures; Statistical Learning; Theoretical Approaches

Further Readings

Doya, K., Ishii, S., Pouget, A., & Rao, R. P. N. (2007). *Bayesian brain: Probabilistic approaches to neural coding*. Cambridge: MIT Press.

Ernst, M. O., & Banks, M.S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429–433.

Kersten, D., Mamassian, P., & Yuille, A. (2005). Object perception as Bayesian inference. *Annual Review of Psychology, 55,* 271–304.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences, 27,* 712–719.

Kording, K. (2007). Decision theory: What "should" the nervous system do? *Science, 318,* 606–610.

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience, 9,* 1432–1438.

Rao, R. P. N., Olshausen, B. A., & Lewicki, M. S. (2002). *Probabilistic models of the brain: Perception and neural function.* Cambridge: MIT Press.

BINDING PROBLEM

Objects have different features such as color, shape, sound, and smell. Some, such as color and sound, are represented separately from the instant they hit our sensory receptors. Other features, such as color and shape, are initially encoded together but subsequently analyzed by separate areas of the brain. Despite this separation, in perception the brain must represent which features belong to the same object. This is the *binding problem.* Any case of the brain representing as associated two features or stimuli that are initially represented separately can be called binding, but this entry focuses on a subset of these: the pairing of features that belong to a common object.

Solutions to the Spatial Binding Problem

A simple solution to the binding problem is to have a single neuron (or other representational unit) for each possible combination of features. However, considering that different feature dimensions such as color, shape, and texture may each have hundreds of values, it is impractical to dedicate a unit to each combination. Still, the visual system does contain neurons selective for certain combinations of features, and these may suffice to solve the binding problem in certain cases.

Wolf Singer has championed the theory that binding is represented via synchronous rhythmic firing of the neurons selective for the paired features. The idea is that the joint activity of the feature representations allows other brain areas to process the features together, excluding features belonging to other objects. Groups of neurons in many parts of the brain frequently do synchronize their responses, and attention to visual stimuli can enhance the effect, but the precise relationship of the phenomenon to perceptual binding remains unclear.

For experimental psychology researchers, two papers on binding by Anne Treisman in the 1980s set the course for nearly two decades. Treisman's feature integration theory (FIT) became the most influential theory of binding and the most influential theory of attention. FIT posits that binding is accomplished by an act of selective attention and has three elements.

1. Features such as color and shape are represented separately in the brain, but each feature (such as red) has a feature map indicating the location of each instance of the feature in the visual field. By virtue of the position of the units that represent the feature relative to the others in the map, a location tag is implicitly included and

activity of a unit signals both feature identity and location.

2. The objects of a visual scene initially cause representations of the various features of all the objects to become active. The system does not yet represent which features belong to the same object.
3. Binding happens when attention is directed to a particular location. The neurons corresponding to this location in each feature map become active, excluding those in other locations, and the features occupying the location are bound.

Spatial Binding: Evidence

The results of thousands of visual search experiments have been interpreted in the framework of FIT. In one case, people were tested on the time needed to find a red "X" target stimulus among a large array of red "O"s and blue "X"s. Because the target is defined by the combination of color and shape, the task requires binding of these features. If there were no binding to determine whether any instances of red were in the same location as the X shape, the target could not be found. Searches that require binding are usually more time-consuming than are searches for a target containing a feature different from those of all the other items. In this latter case, a feature map alone is sufficient to solve the task. According to FIT, the reason for the greater difficulty in the binding search is that the binding step takes time and can only be performed in one location at a time. Attention is hypothesized to visit each location in turn, binding its features until the target is found.

FIT has been useful for explaining visual search results; however, a number of other models, models without any role for attention in binding, can also explain visual search performance. These alternative models imply that the nature of the binding process cannot be determined from visual search results alone. Visual search results are affected by many factors such as image segmentation mechanisms, local salience processing, and crowding, making it difficult to isolate the binding process.

In experiments relying on fewer assumptions, Treisman documented clear binding mistakes with simple displays, in a phenomenon she called *illusory conjunctions*. In one illusory conjunctions experiment, two black digits and three colored letters were flashed briefly on the screen. Participants attempted to report the digits and the letters and the letters' associated colors. They usually reported the correct letters and colors, but occasionally with the wrong pairing. Treisman suggested that the short presentation duration, in combination with the attentional demands required to also report the digits, meant that the attention available was sometimes inadequate to correctly bind the features. Subsequent work verified that illusory conjunctions are perceptual errors rather than an artifact of guessing strategies, and illusory conjunctions are more common when perceptual demands are high, but it remains unclear whether this reflects a critical role for attention or just that binding requires additional sensory processing beyond that required for identifying the features.

To probe the role of attention more directly, Jochen Braun and colleagues devised a quantitative measure of the attentional resources shared by two tasks. Results from this paradigm indicate that contrary to FIT, some feature bindings are perceived at little or no cost to attentionally demanding visual tasks. For example, accuracy in reporting the orientation and color of two line segments in the periphery was almost completely unaffected by the demands of a concurrent task of searching for a T among L's or an L among T's. Results from Braun's paradigm do, however, support the idea that attention is critical for linking features to particular spatial locations. Ability to judge the spatial configuration of adjacent red and green patches (whether red is left of green or right of green) traded off linearly with performance in concurrent central tasks, supporting a critical role for attention. Interestingly, other discriminations that may also rely on configuration do not show this property, for instance, discrimination of face gender or identity. Apparently, the binding required for certain discriminations can proceed with little to no attention. However, even when full attention and extensive processing time is available, binding can fail profoundly, as described in the next section.

Location Tagging of Features: A Prerequisite for Binding?

In her feature integration theory, Treisman suggested that bindings between features are mediated by the features' links to a location in common. Psychophysical demonstrations of binding failures under conditions of full attention provide support for this idea that binding is accomplished through common location tags.

In a display devised by Steven Shevell and colleagues, two separate objects are presented at the same location but in different eyes. For example, a vertical array of orange and gray stripes is presented in one eye, and a horizontal array of gray and blue stripes in the other eye. The conflict between the eyes causes different perceptions to occur alternately, including some involving misbinding of the color and form features. Specifically, arrays of orange with blue stripes of either orientation are perceived. This phenomenon suggests that pairing of features is not fully resolved until after the representations of the two eyes come together. That presentation of multiple instances of a feature in a single location can confound binding supports the notion that binding relies on location tags.

Even when features are perceived in their correct locations, binding can still fail profoundly. Still, this too may reflect a location tagging failure that arises when the spatial scale of analysis of one feature is larger than that of another, allowing the location tag of the larger feature to correspond to more than one instance of the smaller feature. In a display devised by Hugh Wilson and Frances Wilkinson, a pair of dots defines a local orientation. Hundreds of such pairs of dots are scattered across the screen and oriented such that globally, a spiral is perceived. If all these dots are white and unrelated black dots are randomly interspersed with them, then under brief presentation conditions binding fails—people are unable to say whether the dots forming the clearly visible shape are black or white. This phenomenon suggests that the mechanisms for a feature (here, shape) extracted by combining a number of local components do not preserve information about other aspects of its constituents (here, color). Rather, the global shape is assigned to a large area, with multiple colors tagged with locations within it.

Rare neuropsychological syndromes further highlight the role of location tagging. Some patients with bilateral parietal damage erroneously pair color and shape much more frequently than they misperceive the constituent features. Damage to the pulvinar, a subcortical structure, can cause similar binding problems. Both disorders are accompanied by a deficit for localization of even an isolated feature, consistent with the thesis that binding of visual features is accomplished through common location tags.

Temporal Binding

Certain binding failures documented in healthy observers suggest that the binding process is quite slow relative to feature identification. Two dot patterns—each forming a different global shape (such as those mentioned earlier) and each a single color (e.g., red or green)—were set in alternation. The two shapes were constructed such that their shapes could not be determined when the alternation rate exceeded the temporal resolution of the shape identification mechanism. The two global shapes were easily perceived at 15 hertz (Hz), implying rapid grouping of the dot pairs and extraction of the global shape, and the color of each dot was perceptually obvious. However, without extensive scrutiny, observers were unable to determine which shape was formed of green dots and which shape formed of red dots unless the patterns alternated slower than 3 Hz. Binding of color and motion also shows a slow limit of less than 3 Hz, even when both features are local. The alternation of colors and motions causes the display to contain more than one feature of each type in a single perceptual location. With binding based on location tags, this yields binding ambiguity when the displays are alternated fast enough to exceed the temporal resolution of the binding process. The 3-Hz result indicates that binding is slow and requires much more time than does identification of the constituent features.

Even at slow rates, binding of the features that occur together in time may pose problems beyond those faced by spatial binding. For spatial binding, linking features based on common location works well thanks to the large number of spatially organized areas in the visual system. In contrast, no chronotopically organized visual areas have been found, raising the issue of how features might be tagged temporally. Temporal tagging would be unnecessary if all features were processed in the same amount of time. But features have different sensory latencies and processing times. Perhaps the perceptual systems have a scheme for tagging the time that features actually occurred in the world, as opposed to when they are identified by the brain, but this is not yet understood.

From Perception to Memory and Cognition

Once features are bound by perception, these bindings must be maintained by visual short-term memory if we are to do such simple tasks as detect changes in scenes. But visual short-term memory has a limited capacity, and for moving objects, the pairing of object to features may easily be lost. Jun Saiki has used a display with a triangle, circle, and square of different colors moving about the screen and briefly disappearing behind occluders. Sometimes the objects switch colors or shapes while they are behind the occluders, but people are poor at detecting this, even though they do well at detecting a new color or a new shape. Understanding the role of bound object representations versus individual feature representations in real-world behavior requires more investigation of the demands binding places on initial perception and on sustained attention and memory.

Alex O. Holcombe

See also Attention: Divided; Attention: Physiological; Attention: Selective; Visual Search

Further Readings

Braun, J., Koch, C., Lee, D. K., & Itti, L. (2001). Perceptual consequences of multilevel selection. In *Visual attention and cortical circuits* (pp. 215–241). Cambridge: MIT Press.

Clifford, C. W. G., Holcombe, A. O., & Pearson, J. (2004). Rapid global form binding with loss of associated colors. *Journal of Vision, 4,* 1090–1101.

Hong, S. W., & Shevell, S. K. (2006). Resolution of binocular rivalry: Perceptual misbinding of color. *Visual Neuroscience, 23,* 561–566.

Roskies, A. L. (1999). The binding problem. *Neuron, 24* [special issue on binding], 7–9.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 97–136.

BINOCULAR VISION AND STEREOPSIS

Animals such as rabbits and horses have two eyes, one on each side of their heads (binocular vision) that provide them a panoramic view and awareness of the world all around them. For instance, rabbits are able to see potential predators approaching from behind because of the lateral position of their eyes. Humans and a number of other species have forward-facing eyes that do not allow for a view behind the head. What evolutionary advantage could make up for the sacrifice in field of view? One enormous advantage is that the two forward-facing eyes receive images of the same portion of the visual world, but the left and right eye images differ because the eyes view from slightly different positions. These image differences on the two retinas (called binocular disparity) are used by the brain to provide a precise estimate of the depth of objects in the world, relative to where the eyes are looking. This transformation of two images into the perception of depth is called *stereopsis* (Greek for solid sight), and for animals that seek camouflaged prey that are able to remain still, stereopsis can be essential.

Modern humans do not typically use binocular disparity for hunting prey; however, they do rely on this depth cue for daily tasks such as navigating stairs, reaching for the handle of a coffee cup, or threading a needle. Without stereopsis, such tasks become surprisingly difficult; it is only through time and practice that someone who loses stereopsis late in life becomes adept and confident by relying on other, monocular, cues to depth. In conversations with individuals who do not have stereopsis, a common complaint is that, even though long periods have elapsed since they lost stereopsis, they find it difficult to catch a baseball, or they might feel slightly unstable going down stairs. Such losses are experienced by people who once had stereopsis, but then are forced to function without it because of sudden illness or injury. In contrast, many individuals who are born without stereopsis are able to rely on other cues, such as motion parallax, to perform daily tasks and even play competitive sports such as cricket. Unfortunately, others suffer permanent visual deficits, such as amblyopia (often called lazy eye), resulting from the early disruption of their binocular vision.

Although stereopsis is one of many potential cues to depth used by the brain, it is the only one that, on its own and without head movement, provides such a compelling sense of the volume or space between points in the world. It is not known whether individuals who lack stereopsis experience the space between objects in the same way as people with stereopsis do. In a recent case study, Oliver Sacks recounts the story of a woman named Sue who regained stereopsis while in her 40s. She describes the difference in her world, and in her experience of visual space. In one instance, Sue describes eating a meal, looking at her fork and suddenly, for the first time, being aware of the space between the fork and the food she was eating. Her description of this and other experiences of the space between objects and surfaces provides rare insight into a world devoid of stereoscopic depth.

Although scientists and philosophers have studied binocular vision since the time of Euclid (ca. 300 B.C.), it is remarkable that the link between the percept of depth from stereopsis and the two slightly offset images received by each eye was only realized about 150 years ago. For centuries, it has been known that the left and right eyes' images differ but natural philosophers considered it a challenge to explain how the visual system produced an impression of a single world from them. Sir Charles Wheatstone discovered that these differences allow for the perception of a single *three-dimensional* world. He invented a "stereoscope" that used mirrors to present separate images to the two eyes and provided simple stereograms that produced compelling depth when viewed in this way. Wheatstone's 1838 presentation to the Royal Society of London was the starting point for the empirical study of stereoscopic vision. Not only did he give scientists a way to present stereoscopic images to the two eyes, but in doing so, he provided a means of carefully manipulating the input to the binocular visual system. Since then, much research has been focused on the nature of stereopsis, and what stimulus attributes influence this process. His stereoscope was also one of the most successful consumer products of the 19th century and became a fixture in Victorian parlors. The new field of photography provided exciting stereoscopic photographs of natural wonders, exotic places, and the spectacles of the age.

This entry describes the information used for stereopsis, measuring stereoscopic depth perception, and the neurophysiology of stereopsis.

The Information Used for Stereopsis

Stereopsis is based on relatively simple geometry. Figure 1 illustrates this geometry and the information that forms the signal for stereopsis, binocular disparity.

In Figure 1, the observer is fixating object F, and there is a separation between it and another object N. Keep in mind that when we look directly at an object we are positioning the image of that object on the high-resolution portion of the retina, the fovea. The fovea is defined as the reference for a set of geometrically corresponding points (i.e., if you were to slide one retina on top of the other, these points would align) and is itself a corresponding point. The other point, N, is closer to the observer than F is, and the image of this object falls on non-corresponding points: In the left eye, it is to the left of F, and in the right eye, it is to the right of F. The difference in distance between the position of the image of the object and the fovea in each eye is its retinal disparity, which can be represented either as the distance on the retina or in geometric terms as the difference in angular separation of F and N between the two eyes. Because binocular disparity is always relative to fixation, the computation of disparity is the same whether the object is in front of, or behind, the fixation point.

Measuring Stereoscopic Depth Perception

Wheatstone's invention provided a means of presenting and carefully controlling stereoscopic images, while holding other cues to relative depth constant. Much of the research in the 20th century used simple line and dot stimuli to investigate the limits of stereopsis and the factors that determine these limits (see Figure 2, second row). In 1960, Bela Julesz, who was seeking a perfectly camouflaged stereoscopic stimulus, introduced what he called a random-dot stereogram (RDS). He showed that if identical patterns of random dots were presented to the two eyes, and a region in one of the

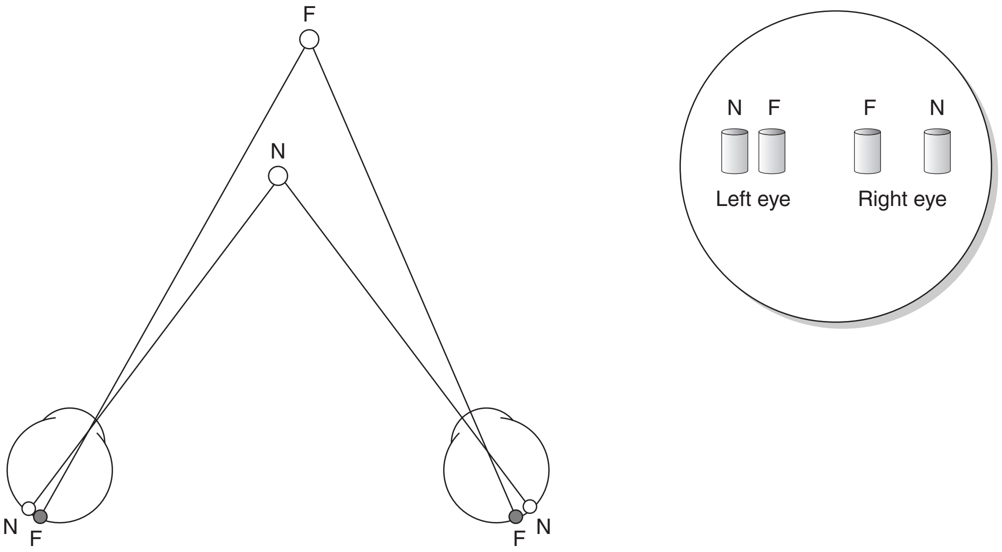


Figure 1 Corresponding and Noncorresponding Points

Notes: The two eyes fixate on object F, and object N sits closer to the observer. As shown in the inset, and by the circles indicating where each object lies on the retina in each eye, the images of N stimulate non-corresponding retinal points. Also, as shown in the inset, the separation between N and F is larger in the right eye because N is shifted away from the midline, toward the left eye.

pairs was shifted laterally (to create retinal disparity), then that region would appear to float above or recede behind the background. Critically, the shape of the shifted region was not visible in either of the monocular views; it could only be seen in the combined or cyclopean view (Figure 2, top row). Further, the amount of depth perceived was proportional to the size of the shift. He argued that this demonstration formed a definitive rebuttal to Gestalt psychologists who believed that the binocular combination required for stereopsis occurred after prior neural processing which extracted form from an image.

Julesz's introduction of the RDS initiated a shift in focus of the experimental community from the relationship between binocular disparity and perceived depth to an issue called the correspondence problem. As argued by Julesz, RDS contain many identical elements, and no obvious contours or form to match in the two eyes. How does the stereoscopic system determine the appropriate disparity given that there are so many possible matches? Julesz proposed that the visual system identifies regions or clusters of elements in two images that are highly correlated, and then extracts a disparity estimate from their relative positions in the two eyes. This cross-correlation approach formed the basis for many subsequent models of stereopsis.

Thresholds and Limits

Under ideal conditions, with high contrast sharp edges and lines, some observers can see the depth provided by disparities as small as 2 to 6 seconds of arc. These disparities are much smaller than that of the spacing of photoreceptors on the retinas and stereopsis has therefore been referred to as a hyperacuity. At a distance of 57 centimeters (cm), such disparities allow discrimination of the depth of objects separated by less than the width of a typical human hair! Such stereoscopic thresholds, or stereoacuities, reveal the extraordinary ability of the stereoscopic system under ideal conditions, and with some practice. Experiments in which large numbers of young adults, without additional training, have been assessed have shown that 80% of those tested can reliably discriminate disparities as small as 30 seconds of arc, which is less than the thickness of a fingernail at arm's length. Such fine stereoacuity judgments depend on a number of stimulus attributes such as visibility, viewing time, and blur.

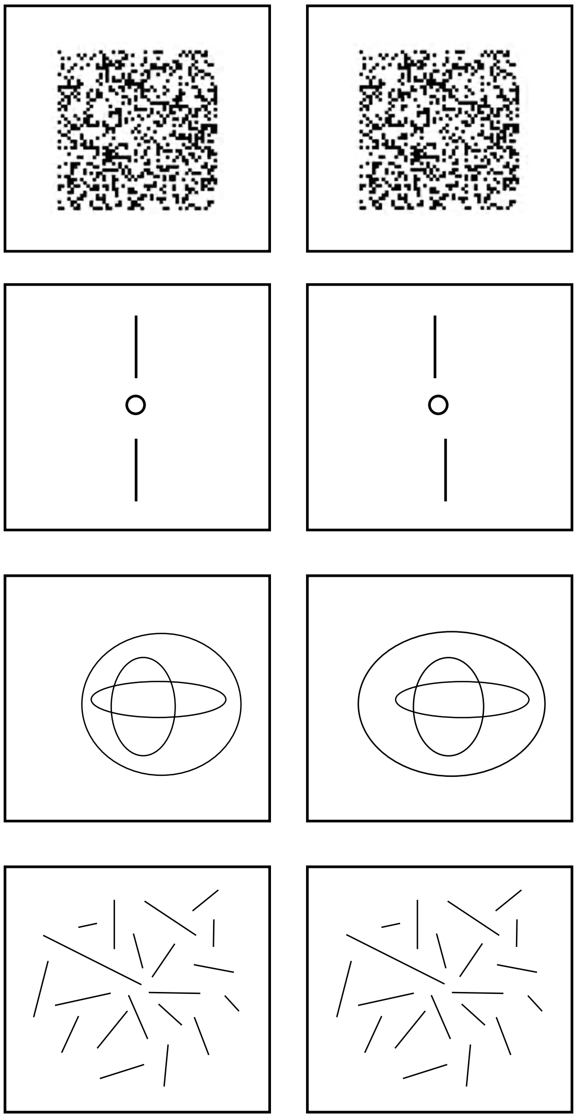


Figure 2 Stereograms

Notes: The stereo pairs shown here illustrate how space and form can be created simply from lateral shifts in images presented to the two eyes. The easiest way to view stereograms is to put them in a stereoscope, which presents the two images separately to each eye. You can also view these without an optical device by first crossing your eyes, and as the images become doubled in the two eyes, allow the central pair to overlap. Then, fixate on this central fused image. Gradually it will sharpen, and you will be able to see the three-dimensional (3-D) relief in the figures. In the first row is a random-dot stereogram, like that created by Julesz, which contains a central square protruding in depth. Row two contains a simple line stereopair; the top bar should recede in depth and the lower bar will come forward. Row three shows circles at different positions in depth. In row four, two of the line elements are displaced in depth, one near and the other far. (Note that these perceptions may be difficult to achieve without a stereoscope.)

The stereoscopic system, although useful for making high-resolution depth judgments, also provides relative depth information for stimuli that have much larger disparities. In fact, the binocular disparity can be so large that the images of the objects in the two eyes cannot be fused, and are seen as double, but still the observer can reliably judge the position in depth of the object relative to fixation. Investigators have shown that the upper limit for stereopsis, or the largest disparity that produces a percept of depth, generally occurs when the binocular disparity is near three times the width of the object. Thus, the upper limit scales with object size. However, this limit, unlike stereoacuity, is not affected by the amount of high-contrast detail within the object, stimulus contrast, or the viewing time.

Development of Stereoacuity

Studies of stereoacuity in infants in the 1980s used a variety of stimuli (RDS and line stereograms) and methodologies (e.g., behavioral measures and visual evoked potentials) to show that stereoscopic depth percepts do not reliably appear until approximately 4 months of age. Evidence suggests there is a critical developmental period between 4 and 24 months of life. If normal binocular vision is prevented by disease or visual defects during this time then development of stereopsis is highly improbable in later life. From this age on, it appears that stereoacuity continues to improve until 9 to 12 years, at which point performance reaches adult levels. However, some studies suggest that further improvements in stereoscopic vision occur with age until 21 years.

The Neurophysiology of Stereopsis

Electrophysiology

At about the time that Julesz was developing the RDS, David Hubel and Torsten Wiesel were conducting groundbreaking research, recording the responses of binocular neurons in the cat visual cortex. Their Nobel-prize winning research identified and classified cells in the primary visual cortex (known as V1 or striate cortex in primates) according to eye dominance. They found that a large portion of the cells received input from both eyes.

These binocular neurons were subsequently shown by Horace Barlow and colleagues to be disparity tuned, responding to specific retinal disparities between objects on the two retinas. Subsequent experiments with awake monkeys suggested that disparity-selective neurons could be classified as tuned excitatory, tuned inhibitory, near or far. The tuned neurons respond to small disparities, near fixation, whereas the near and far cells respond to coarser disparities in front of and behind fixation, respectively. More recent experiments show that these are not strict categories, but that the cell types fall along a continuum.

Electrophysiological studies have shown that V1 cells respond to anti-correlated RDS in addition to the traditional correlated RDS described. In anti-correlated RDS, the polarity of the elements (black/white) is reversed in the two eyes so that a white element in the left eye would correspond to a black element in the left eye and vice versa. The response of V1 cells to anti-correlated RDS suggests that these units do not code depth percepts per se (humans do not see reliable depth via disparity in these anti-correlated patterns). However, neurons in area V2 do distinguish between these types of stimuli, and respond to correlated patterns in a manner that is consistent with behavioral reports. Other studies of the disparity selectivity of extrastriate neurons have shown disparity processing in areas thought to be dedicated to other visual attributes, such as motion in the middle temporal (MT) cortex. Investigators have influenced monkeys' behavioral responses to the location in depth of a stimulus by applying a mild current to MT neurons. In these experiments, only coarse disparity tuned neurons were influenced by the microstimulation, suggesting that there is a distinction between coarse and fine disparity processing at this level in the system. It has been proposed that the dorsal pathway, which includes MT, is specialized for processing coarse disparities and driving eye movements. According to this description, the ventral stream, which terminates in the inferotemporal cortex (IT), is responsible for fine disparity and high-resolution depth judgments.

fMRI Studies

The study of the neural basis for stereopsis in human subjects has grown rapidly with the advancement of brain scanning technology such as functional magnetic resonance imaging (fMRI). Investigators had previously shown that individuals with damage to the parietal region suffered deficits in stereoscopic depth perception, and subsequent studies using fMRI confirmed this relationship. However, more recent work by Ben Backus and colleagues demonstrates that, as electrophysiology has shown in the monkey cortex, processing of binocular disparity information occurs throughout the human visual cortex. To date, the results of fMRI studies in humans map well onto the electrophysiological results. Particularly interesting is recent work that suggests that MT is involved in the combination of disparity and motion for the perception of 3D form.

Although researchers are beginning to understand the neural and behavioral basis of stereopsis, many questions remain. For instance, is stereopsis processed differently in the dorsal and ventral streams, is surface perception a critical component of stereoscopic processing, and can stereopsis be regained by individuals previously believed to be stereoblind?

Laurie M. Wilcox and Robert S. Allison

See also Animal Depth Perception; Depth Perception in Pictures/Film; Spatial Layout Perception, Psychophysical

Further Readings

Arditi, A. (1986). Binocular vision. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance.* New York: Wiley.

DeAngelis, G. C., Cumming, B. G., & Newsome, W. T. (1998). Cortical area MT and the perception of stereoscopic depth. *Nature, 394,* 677–680.

Howard, I. P., & Rogers, B. J. (2002). *Seeing in depth* (Vol. 2). Toronto: I. Porteous Press.

Julesz, B. (2006). *Foundations of cyclopean perception.* Cambridge: MIT Press. (Original published in 1971)

Regan, D. M. (Ed.). (1991). *Binocular vision (vision and visual dysfunction)* (Vol. 9). Boca-Raton, FL: CRC Press.

Sacks, O. (2006, June 19). Stereo Sue. *New Yorker,* 64–73.

Von Noorden G., & Campos, E. (2002). *Binocular vision and ocular motility.* St. Louis, MO: Mosby.

BISTABLE PERCEPTION

Human perception is amazingly efficient and usually stable despite the intrinsic ambiguity and uncertainty of sensory input. However, occasionally our sensory systems do fail to reach a single interpretation for a given input. Sometimes, our percepts fluctuate between two distinct interpretations or states, although more than two interpretations are also possible. This is called *bistable* (or multistable for more than two states) *perception.* The perceptual alternations are spontaneous and stochastic, and the alternating percepts are often mutually exclusive. This entry describes varieties of bistable perception, binocular rivalry, mechanisms for perceptual switching, and tools used for studying mechanisms of object perception and visual awareness.

Varieties

Several varieties of stimuli can lead to bistable perception. Many bistable phenomena are generated from a set of common factors, some of which are listed here.

Unspecified Depth Relationship

Retinal images are fundamentally two-dimensional, but objects in the real world are three-dimensional, thus the depth relationships between different parts of an image are ambiguous. The Necker cube—a wireframe cube that can be seen as switching back and forth between two different viewing angles—falls into this category. The depth of the different wire segments is intrinsically unspecified.

Unspecified depth information can also lead to bistable perception of dynamic images. For example, varying speed of motion across different parts of an image can lead to the perception of a three-dimensional structure (structure-from-motion, motion parallax cue). Two sets of random dots sliding sideways across each other within a rectangular area, their speeds following a sinusoidal function, slower near the boundary and faster in the center, would be seen as a transparent rotating cylinder. However, the depth relationship of the two surfaces is ambiguous and bistable. Observers can see either the leftward moving dots or the rightward moving dots as the front surface, and as a consequence, perception of the cylinder alternates between clockwise and counter-clockwise rotation motion. A rotating silhouette figure is also bistable and can be perceived as rotating either clockwise or counter-clockwise. This is also because the depicted person could have two possible viewpoints (facing direction) at any moment, essentially also a problem of unspecified depth relationship.

Ambiguous Border Ownership

Boundaries between different regions of an image can have ambiguous ownership. The familiar Rubin's face-vase illusion belongs to this category. Here the space between a pair of facing profile faces can be seen as a vase. When the percept of the vase arises, the faces recede, and vice versa. This perceptual alternation occurs automatically. In this case, the borders between the two outer faces and the middle vase can be seen as belonging to the faces (face percept) or belonging to the vase (vase percept). In essence, the ambiguity of border ownership is also an ambiguity in depth assignment—which surface area is in front and occludes the other regions?

Competing Grouping Possibilities Between Image Components

Image components can be grouped in different ways. A simple example is a matrix of evenly spaced dots that can be seen either as rows of dots or columns of dots. The bistability can be enhanced when different Gestalt cues lead to different groupings (e.g., color alternating across rows and size alternating across columns). Similarly, moving components can be grouped in different ways, resulting in bistable (or multistable) percepts, as exemplified by the motion quartet demonstration, in which case grouping (or correspondence) occurs over time.

More generally, bistability can result from components fluctuating between organized and unorganized states. A good example is the well-researched component versus pattern motion: two gratings sliding across each other can be seen as just that, or they can stick to each other and form a plaid. Under the right conditions, perception clearly alternates between component grating motion and grouped plaid motion. Another excellent example is a moving diamond placed behind three invisible occluding bars so that no corner can be seen. The four visible edges of the diamond can be seen as moving relatively independent of each other (sliding up and down in the space between occluders), or they could be seen as parts of a single diamond shifting sideways.

Bistable Perception in Audition and Tactile Perception

Although less common, bistable perception also occurs in other sensory modalities. For example, bistable streaming can occur with two repetitive tones when the frequency difference between them is moderate. Over time, perception alternates between one combined stream and two independent streams. Similar to apparent motion experienced visually, alternate tactile stimulation at two nearby locations generates a tactile apparent motion. Interestingly, a tactile version of the bistable motion quartet can be experienced as well. For example, if pairs of brief vibrotactile stimuli are applied to the finger pad with short intervals, and with each successive stimulus pair alternate between the opposing diagonal corners of a small square, observers will experience switches between the two motion paths, one that moves up and down and another left and right.

Binocular Rivalry

A special yet intriguing type of bistable perceptual phenomenon is binocular rivalry. Normally, our two eyes' retinal images are similar. However, when the images at corresponding retinal locations of the two eyes are different, observers tend to see an alternation between these two images, a phenomenon called binocular rivalry. For example, if a horizontal grating and a vertical grating are projected to the left and right eyes' corresponding locations, respectively, observers would see an alternation between the horizontal and the vertical grating.

One of the central questions with regard to binocular rivalry has been "what rivals during binocular rivalry?" Two main answers have been provided to this question. (1) Rivalry is a mutual inhibition between the two sets of eye-specific neurons (eye rivalry); (2) rivalry is an alternation between two stimulus representations (stimulus rivalry). Each view has its own strong supporting evidence. More recently, the consensus has been that binocular rivalry involves both early (low level) and late (high level) processes, and both eye rivalry and stimulus rivalry are possible. Whether it is primarily eye rivalry or stimulus rivalry is determined by specific stimulus conditions.

The visual system consists of multiple and parallel pathways, and growing evidence indicates that rivalry occurs in the parvocellular rather than in the magnocellular pathway. For example, in the absence of form and color conflict, two flickers with a dramatic difference in temporal frequency are not engaged in rivalry, and fast motion signals are much less likely to rival than are slow motion signals, even with conflicting directions of motion between the two eyes.

Mechanisms for Perceptual Switching

A hallmark of bistable perception is that one has little voluntary control over the switching process. Attention can modulate the dominance duration of a percept to some extent, more so for certain bistable pictures than for binocular rivalry, but perceptual switching cannot be halted. The dynamics of bistable perception is usually characterized as a stochastic process. The distribution of percept durations has often been described (or fitted) with a gamma distribution; however, alternative models have been suggested, among them that the gamma distribution is better fit to switching rates than to percept durations. The neural mechanisms responsible for perceptual switching during bistable perception remain poorly understood. However, adaptation to the currently dominant percept/stimulus plays an important role. The idea is that over time the dominant percept/stimulus would become weaker and succumb to the previously suppressed percept/stimulus. Reducing the effect of local adaptation (e.g., by slowly moving the pair of competing stimuli) significantly slows binocular rivalry alternations. Another interesting proposal is that bistable perception, including binocular rivalry, results from switching dominance between the two hemispheres. Loosely related to this proposal, rivalry switching rates in certain clinical populations (e.g., bipolar patients) were found to be much slower than were normal controls.

Although perceptual switching cannot be halted voluntarily, manipulating the context (spatial, temporal context, or cross-modal context) of the bistable stimuli can have strong effects on the dominance durations and switching dynamics. For example, placing a Necker cube in the middle of unambiguous cubes consistent with one of the percepts would significantly increase the probability of seeing the Necker cube the same way as those unambiguous ones in the context. Intermittent presentations of bistable stimuli, with stimulus on-times less than the average dominance duration and stimulus off-times a little longer, can lead to a dramatic reduction in switching. In some observers, such a temporal manipulation can even completely abolish the perceptual switching, thus stabilizing the bistable stimuli. With intermittent presentations, the forces (e.g., adaptation) that drive switching from one percept to another are apparently much weaker than the facilitation (e.g., priming of some sort) from prior experience of the dominant stimulus. Unambiguous cues from other sensory modalities can bias the percept of an ambiguous stimulus. For example, touching a rotating cylinder can significantly bias the percept of a visual bistable rotating cylinder so that the two cylinders match rotation direction.

Tools to Study Mechanisms of Object Perception and Visual Awareness

Bistable stimuli have received renewed interest in recent years partly because they afford the opportunity to isolate neural mechanisms more closely related to awareness. The logic is relatively simple: With bistable perception, the stimulus stays constant but perception changes, thus correlated changes in neural activity in the brain must have more to do with the perceptual change. Conversely, binocular rivalry can be used to create a complementary situation where monocular stimuli presented under suppression are not explicitly detected by the observer, and the neural mechanisms registering the unperceived stimulus would less likely serve as the neural mechanisms directly involved in generating perceptual awareness. The observed neural signal change is not necessarily the so-called neural correlates of consciousness (NCC), but we can at least conclude that neural mechanisms with activity more in sync with perceptual states are better candidates for NCC than are neural mechanisms that are more in sync with stimulus states. For the goal of presenting stimuli under interocular suppression, an interesting modification to the rivalry paradigm is the so-called continuous flash suppression paradigm in which one eye is presented with high-contrast dynamic noise and the corresponding retinal region in the other eye is presented with interesting images. Often, the goal in this situation is to check at what level in the processing hierarchy as well as in which pathway the suppressed images can be processed.

Certain bistable stimuli can also be used to study visual neural mechanisms of object perception. As described earlier, a class of bistable stimuli fluctuates between an unorganized collection of components (features) and an organized perceptual pattern (object). Examples include the moving stimulus that could be seen as either two sliding gratings or a single plaid. Such stimuli allowed the discovery of motion sensitive neurons that are selective for direction of component motion and motion neurons selective for the direction of pattern motion. In the case of a diamond moving behind invisible occluding bars, the alternate bistable percepts correspond to either a coherent shape (diamond) or elementary features (moving bars). With this stimulus, an interesting finding is that when observers report seeing the diamond, activity in the early visual cortex decreases while activity in high-level object sensitive areas increases. This observation provides insight into the interactions between the primary visual cortex and object sensitive cortex.

Sheng He

See also Auditory Scene Analysis; Context Effects in Perception; Multimodal Interactions: Visual–Auditory; Unconscious Processes

Further Readings

Blake, R. (2001). A primer on binocular rivalry, including current controversies. *Brain and Mind, 2*(1), 5–38.

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience, 3*(1), 13–21.

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences, 3*(7), 254–264.

BODY PERCEPTION

Body perception refers to the detection and recognition of an animate body and its postures. The body is a special object of perception because unlike other objects, it refers to perception of the self as well as to perception of others. Our own bodies tell us what we are doing now and help us figure out how we should perform actions that give us what we want. Visual perception is integrated with additional sensory channels that provide tactile, proprioceptive, and kinesthetic information for multimodal processing. We see it, feel it, and control it. This online body information organizes current body information with stored spatial and structural body information. Because of the body's unique structure, other people's bodies are also special objects of perception. They give us information about what they are doing and whether we need to respond to what they are doing. We frequently observe other people's bodies, imitate them, interact with them, and predict their movements. Hardwired brain circuits capture specific body configurations, postures, and actions. Finally, body perception involves understanding the relation between the two sources of body information. The influence of one's own body on the perception of other people's bodies defines embodied perception. This interaction of body information from one's own body with information from other people's bodies helps us understand how other people are like us and helps us learn new skills.

Two types of representations are specific to the body and are not used for other objects: an online representation that coordinates sensorimotor inputs about our current body configuration and a long-term body representation that organizes information about the structure and spatial properties of the human body in general. Although terms such as *body percept, body image,* and *body schema* have been used interchangeably to refer to both of these concepts, the term *body schema* is most frequently used to refer the online representation, and the term *body structural descriptions* is used to refer to the long-term, spatial properties of the body. This entry describes perception of one's own and other people's bodies, specialized body processing, and the significance of body perception.

Perception of One's Own Body

Perception of our own body relies on the integration of information from multiple senses into a common framework. These sensory inputs combine to form an online body representation that tells us about the configuration of our body parts in a particular body position at a certain point in time. This helps provide the information needed to produce successful actions. This dynamic representation changes every time the body changes position. Inputs from the senses and muscles (e.g., touch, kinesthesis, proprioception, and vision) modify the immediate, online body representation and update it to keep track of current body positions and current locations of sensation. Thus, the online body representation represents one's own body exclusively because it contains privileged information about present body status, sensory feedback, and motor intention for performing actions in the environment.

A number of studies have documented the joint contributions from visual, tactile, and proprioceptive inputs to our perceived location and identity of body parts. For example, when a person is presented with the simultaneous presentations of tactile and visual stimulation, her response speed to the *tactile* stimulus is slower when the *visual* stimulus is close to her hand compared with when it is far from her hand. The same pattern of responses occurs even when the viewed arm and hand are replaced with a fake arm and hand in an anatomically possible posture.

Moreover, the relative contributions of vision and touch to body perception can be revealed by introducing sensory conflict. Distorting devices such as prisms, mirrors, and video monitors can be used to artificially displace the visual location of body parts to varying degrees from natural viewing conditions. In other words, the visual image of our hands can be displaced to appear somewhere else in space or time. People experience a conflict between the seen and felt positions of their body parts when their visual fields are distorted or displaced by wearing prism glasses. This exposure to sensory conflict leads to a shift in the perception of the body part so that it "feels" like it is located where it is "seen." These findings indicate the importance of visual contributions to body perception.

The importance of proprioception and kinesthesis to body perception can be demonstrated through a body perception illusion. Vibrating the tendons of arm muscles induces illusory arm movements. For example, it feels as though one's bent arm is straightening if the biceps muscle is vibrated. If the biceps muscle is vibrated while the hand of that arm holds the nose, the experienced sensation is that one's nose "grows" or lengthens like Pinocchio's when he tells a lie. This illusion constitutes a resolution of sensorimotor conflict with the long-term body representation. Because the head and body are stationary, the combined sensory input from the hand and the arm is perceived as the nose elongating. Thus, the configuration of the whole body is inferred from the various inputs rather than from the direct perception of the body parts. Our structural body representations interact with online representations even without vision.

The perception of one's own body in some ways is different from the perception of other people's bodies. For example, these perceptions differ in visual and proprioceptive inputs. Despite these apparent differences, however, the perception of one's own body and of other bodies is closely related.

Perception of Other People's Bodies

Body representations also assist in the organization and perception of other people's bodies. The interface between the perception of one's own body and the perception of other people's bodies is the long-term, body structural representation. This representation is a dedicated, internal representation of the spatial organization of body parts and the invariant properties of human body structure. The representation incorporates information about the relative locations among body parts, the degrees of freedom for movement at joints, and anatomically possible body function. It is multimodal, incorporating information about one's own and others' bodies from multiple senses, especially vision, touch, and proprioception. In contrast to the online body representation, the long-term body representation is used to represent body information about other people as well as oneself.

One source of evidence for a long-term, multimodal spatial representation of body comes from the adult neuropsychological literature. Patients with autotopagnosia often have brain damage in the left parietal lobe. Although patients can identify individual body parts by themselves, they cannot locate the same body parts within the context of the complete body. Patients cannot localize body parts on themselves, on the experimenters, or on dolls. This deficit is not limited to vision because patients cannot find specific body parts on bodies whether they look for them or feel for them. Further, it is specific to the body because patients can find named parts on other objects such as trucks or trains. Finally, the deficit affects the spatial body representation specifically because patients can demonstrate good knowledge of bodies, naming of body parts and relatively intact general spatial abilities. Patients with autotopagnosia appear to have lost their spatial body representation that is used for organizing information from their own bodies and other people's bodies.

The structural similarity between one's own body and other people's bodies helps the visual system make inferences about other people and their actions. From these similarities, it can determine whether one's own body could perform movements similar to those of someone else. For example, people spontaneously use their own body to make judgments about other bodies. In one study, people were given pictures of bodies rotated to various degrees away from vertical and were asked whether a dot was on their left or right hand. Response times varied depending on the discrepancy between the angle of the illustrated hand and the person's own hand. The response times indicated that people mentally simulated the movement of their own hand to match the picture before making the left or right hand judgment.

The use of one's own body to interpret the visual world is especially important for perceiving other human and animate bodies and their actions from impoverished visual information. The phenomenon of biological motion refers to our ability to readily perceive a solid human or animal walking and performing complicated actions such as running, yoga, or dancing from moving dots or point-light displays. Point-light displays are created from filming a person moving in the dark with lights attached to his or her joints. Only the moving lights against a dark background are visible; no explicit body structure information is provided in the stimulus. Even with such minimal body information, people can identify human and animal body structure and movement almost immediately. Although humans can perceive the biological motion of animals, they are most efficient at recognizing other humans. They can even distinguish themselves and friends from strangers. It is proposed that people use their representations of their own bodies and visual experiences to interpret the displays.

Other people's motions influence body perception, but our own movement influences body perception as well. What we do with our own bodies affects how we perceive the bodies of others. In one study, people performed a memory task in which they compared two sequentially presented postures of another person. At the same time, they either moved their arms or their legs. Performance in the posture memory task depended on what limbs were moved. If participants moved their arms, their memory for other people's arm postures was selectively improved. When participants moved their legs, their memory for other people's leg postures was selectively improved. Importantly, the interaction between body movement and memory is specific to the body. When people compare two nonsense objects (e.g., made from LEGO toys), movement has no effect on memory. Thus, the same body representation appears to be used to encode the body positions for the self and others.

The integration of one's own body information with other people's body information is important for learning new skills and social interactions. We learn from other people through imitation. Imitation unites the perception of another human's action with one's own production of that action. For example, when learning a tennis serve, a person perceives the action visually, and reproduces the action motorically. The commonalities between what people can do and what they perceive other people doing suggest that a similar mental code and neural system support the learning of new skills and understanding what other people are doing. The mirror neuron system has been proposed as a neural network that responds similarly when a primate observes the execution of a goal-directed action (e.g., reaching for a peanut) and when the primate physically executes the action. Ultimately, these systems help organize body information into meaningful perceptions that tell us how to prepare ourselves for responding appropriately to what others are doing.

Specialized Body Processing

In addition to having body-specific representations, human bodies and their postures are visually perceived using "special" recognition processes that are not used by most common objects. Like faces, bodies are recognized using configural processing mechanisms. There are good reasons why bodies are perceived more as faces are than like other objects. For example, both faces and bodies are important social stimuli for which recognition has evolutionary importance. Humans have extensive experience viewing faces and bodies from birth. Both are identified at the individual level (e.g., Joe) rather than as just a category (e.g., face or body). Importantly, humans can configure their faces and body postures to convey emotions, intentions, and other meaningful information.

The visual perception of body postures (and faces) depends on configural processing rather than on the part-based processing used to recognize most objects. To recognize a body posture, one cannot rely on the presence or absence of a particular part. Bodies usually have two arms, two legs, a head, and a trunk. Instead, different body postures are identified by differences in the precise spatial relations among the shared object parts, or their configuration. In contrast, part-based recognition processes are used for most common objects. Recognition time and accuracy are largely independent of where parts are located or in what orientation the object is situated. A widely used empirical indicator of configural processing is the inversion effect: the recognition of upright items is faster and more accurate than is the recognition of inverted items. Inversion disrupts the configuration or spatial relations among features. For bodies, inversion disrupts the basic body-part relations because heads are no longer above the trunk, which is no longer above the legs. Robust inversion effects have been found for both body postures and faces, suggesting that they are processed configurally. In contrast, most objects (e.g., houses) do not show inversion effects.

The inversion effect can also be used to indicate what body information is necessary to define the body. Researchers manipulated the body stimulus to find out what changes in body information lead to the breakdown of configural processing as indicated by the loss of the inversion effect. When bodies were divided in half along the vertical axis (dividing left from right), the inversion effect was robust. It appears that the visual system was able to reconstruct the body from the long-term body representation because the body is largely symmetrical along its vertical axis. However, when bodies were divided in half along the horizontal axis (at the waist), the inversion effect disappeared. Individual body parts (e.g., arm, leg, head) did not produce inversion effects either. Of interest, when body-part relations were scrambled and limbs were placed on the trunk in alternate locations (i.e., arms put in the head position, legs in arms, etc), not only were no inversion effects found, but the visual system's ability to distinguish between scrambled body positions plummeted. In other words, when the spatial locations of limbs on the trunk violate the long-term spatial body representation, body posture perception is severely impaired.

In addition to functional processing mechanisms, the brain provides a special status to perceptual stimuli that correspond to the human body. Neuronal populations in the superior temporal sulcus and fusiform gyrus are selectively sensitive to bodies compared with other objects and faces. One neuroimaging study compared neural responses to images of headless bodies and body parts to control images. A region of the lateral occipitotemporal cortex, known as the extrastriate body area or EBA, responded selectively to static images of human bodies and body parts. In contrast, the EBA had only a weak response to face, object, and object parts. The response of the EBA also generalized to line drawings, stick figures, and silhouettes of bodies. It appears that the EBA represents bodies at an abstract level that does not rely on visual features.

Thus, the way we view the human body is distinctive from the way we view most other objects. Behavioral and neural evidence suggests the body appears to be defined by its hierarchical structure and that the specific configuration of body parts is important for recognition.

Development of Body Perception

Developmental researchers have investigated what aspects of human bodies allow infants to distinguish them from other objects. A sensitivity to human bodies appears early in development. Evidence from phantom limb patients points to the existence of primitive spatial body representation that may be hardwired in the brain. Despite a lack of sensory input and experience using limbs, individuals with aplasia (e.g., born without limbs) often experience phantom limb sensations and perceive movements of other people's limbs no differently from typically developing individuals.

In typically developing children, a sensitivity to the configuration of body parts appears sometime after the first year of life. Infants younger than 18 months are unable to distinguish between normal and scrambled images of bodies. Although infants' perceptual expectations about normal human faces develop earlier than do those about human bodies, a human body template is shared among infants.

During the first year of life, infants also acquire a sensitivity for the biomechanics of how humans move within the organization of the human body. Three-month-old infants prefer point-light walker displays of upright walking humans relative to inverted walking humans. By 5 months, they are able to distinguish a global body form as long as it provides a context in which the body parts are salient. Infants' visual sensitivity to the structure of the human body corresponds with the infants' acquisition of a basic set of motor skills during the first 18 months. Thus, the development of sensitivity for the visual perception of human bodies corresponds with the development of motor production.

Researchers have proposed that infants, like adults, use body representations to integrate and process stimuli across sensory modalities. Newborns are able to copy gestures. For example, when an adult opens his mouth, the infant opens his mouth in response. When an adult sticks out his tongue, the infant sticks out his tongue. The ability of infants to view another person's movements and reproduce them with their own bodies indicates that infants have body representations that link the infant's actions (proprioception) to the actions of another (vision). These rudimentary imitation skills suggest that neural systems, such as the mirror neuron system, are in place early in development to facilitate learning from other humans.

Significance

Humans are embodied and optimized for processing "like me" information. Perceptual processes integrate information about the observer's own body and other's bodies. Further, visual processing changes if stimulus features resemble information in the observer's own body representations. Specialized body representations and mechanisms help organize information from other people's bodies and aid in body perception. Why is efficient body perception important? The significance of interpersonal action suggests that a key function of these mechanisms is social perception. People's perceptual goals are functionally based around what other people are currently doing; knowing what another is about to do increases the chances that one's own actions will be appropriate.

Catherine L. Reed

See also Body Perception: Disorders; Embodied Perception; Face Perception; Infant Perception; Mirror Neurons; Object Perception; Social Perception; Vision

Further Readings

Maravita, A., Spence, C., & Driver, J. (2003). Multisensory integration and the body schema: Close to hand and within reach. *Current Biology, 13,* R531–539.

Meltzoff, A. N. (2007). "Like me": A foundation for social cognition. *Developmental Science, 10,* 126–134.

Meltzoff, A. N., & Prinz, W. (Eds.). (2002). *The imitative mind: Development, evolution, and brain bases.* Cambridge, UK: Cambridge University Press.

Parsons, L. M. (1987). Imagined spatial transformations of one's hands and feet. *Cognitive Psychology, 19,* 178–241.

Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience, 8,* 636–648.

Pinto, J., & Shiffrar, M. (1999). Subconfigurations of the human form in the perception of biological motion displays. *Acta Psychologica: Special Issue on Object Perception & Memory, 102,* 293–318.

Prinz, W., Shiffrar, M., Thornton, I., Knoblich, G., & Grosjean, M. (Eds.). (2006). *The human body: From the inside out,* Oxford, UK: Oxford University Press.

Reed, C. L., & Farah, M. J. (1995). The psychological reality of the body schema: A test with normal participants. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 334–343.

Reed, C. L., Stone, V., Bozova, S., & Tanaka, J. (2003). The body inversion effect. *Psychological Science, 14,* 302–308.

Shiffrar, M., & Freyd, J. (1990). Apparent motion of the human body. *Psychological Science, 1,* 257–264.

Slaughter, V., & Heron, M. (2004). Origins and early development of human body knowledge. *Monographs of the Society for Research in Child Development, 69,* 1–102.

Stevens, J. A., Fonlupt, P., Shiffrar, M., & Decety, J. (2000). New aspects of motion perception: Selective neural encoding of apparent human movements. *NeuroReport, 11,* 109–115.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review, 9,* 625–636.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin, 131,* 460–473.

BODY PERCEPTION: DISORDERS

The brain computes at least three distinct types of representation of the human body. One type of representation, known as *body semantics* (or *body image* in the classical literature) codes long-term explicit knowledge about body parts, their functions, and their associates (e.g., the wrist is the place where a watch or bracelet is worn). The second type of representation provides a dynamic, constantly updated ("online") mapping of the current locations of body parts in space (the *body schema*). The third kind of representation codes more stable, long-term topological information about the structural positions of body parts on the body surface (*body structural descriptions*). Although the brain structures mediating these representations are not firmly established, several lines of evidence suggest that the body schema and body structural descriptions rely on the integrity of the parietal lobes, whereas body semantics are likely mediated primarily by the temporal lobes. Damage to the body schema or body structural descriptions results in a diverse group of disorders of body perception, to be described in this entry.

Body Schema Disturbances

The body schema is an online representation of the body derived from numerous motor and sensory inputs (e.g., proprioceptive, vestibular, tactile, and visual). It is used to generate accurate representation of body posture, to guide action, and to aid in the perception of the body postures and actions of others. Patients with disorders of the body schema have difficulty performing such tasks.

Personal Neglect and Associated Phenomena

Hemispatial neglect is a common consequence of stroke, particularly to the brain's right hemisphere, but also occurs on occasion following left hemisphere stroke, as well as subsequent to traumatic brain injury or tumor. Patients with neglect fail to attend to or act in the side of space opposite the brain lesion and may fail to attend to one side of their bodies. Thus, for example, patients with left neglect caused by right hemisphere lesions may fail to wash their left hands, comb their hair on the left of their heads, or shave the left sides of their faces. Even when looking in a mirror, they frequently fail to recognize that the left side of the body is poorly attended. This is known as *personal neglect.*

A well-known phenomenon in the neglect syndrome is that of *tactile, auditory, or visual extinction.* It has recently been demonstrated that the extinction phenomenon is observed within single body parts. For example, if a neglect patient is touched simultaneously in two spots on the back of the same hand, one spot slightly to the left of the other, he or she will frequently only report the right-most touch. There is thus a "gradient" of attention that spans individual body parts.

Neglect patients often exhibit *anosognosia* for (i.e., unawareness of) their motor and spatial deficits. For example, patients with weakness of an arm or leg after stroke may frankly deny that there is a problem, even when asked to demonstrate the limb's function. Finally, and less commonly, patients with neglect may exhibit *somatoparaphrenia,* or delusions that the limbs on the side of the body opposite the stroke do not belong to them, and may even belong to others. It is not known precisely how such body-related delusions may emerge, but it is likely that they are the product of a mismatch between intact and damaged parietal lobe sectors that code different aspects of perception and action of the limbs.

These deficits all appear to be different facets of a disturbance in the body schema. Treatments that act on the integration of body and space, such as irrigating the ear on the damaged side with cold water to stimulate the vestibular system, may temporarily ameliorate the deficits. Longer-lasting effects have been difficult to obtain.

Ideomotor Limb Apraxia

Ideomotor apraxia is a disorder common in left-hemisphere stroke and degenerative dementia (e.g., Alzheimer's disease, corticobasal degeneration) that is characterized by spatiotemporal errors in pantomime (e.g., "pretended" use of a tool, without the tool in hand), imitation, or recognition of complex movements and postures. For example, a patient with ideomotor apraxia might imitate a hammering gesture with an increased-frequency, low-amplitude movement in the horizontal plane, with the hand open rather than clenched. Such patients are often also impaired in the imitation of meaningless movements and postures, and some of these patients are unable to position a mannequin to copy the same meaningless poses. This indicates that the problem in imitation of meaningless movements and postures is at least partly a problem in perception, and not just in execution. The performance of patients with ideomotor apraxia frequently improves considerably when they are permitted to actually hold and use a tool rather than simply pantomiming the action.

At least some of the differences between pantomimed and actual tool use may be explained because pantomime relies heavily on a form of spatiomotor coding of limb position, called *intrinsic spatial coding,* which is different from the *extrinsic spatial coding* that is strongly involved in tool-use tasks. In an intrinsic spatial coding system, the positions of the limbs and joints are coded with respect to an internal space centered on the body parts, that is, without reference to the external world.

Intrinsic spatial coding must be used, for example, when positioning the body parts into a desired configuration with eyes closed (for example, try pantomiming a "sawing" motion with eyes closed. The fact that you maintain a fairly fixed set of coordinates during the process indicates that your intrinsic spatial coding is intact). In contrast, *extrinsic spatial coding* systems explicitly code the position of body parts with respect to the external world. The latter is particularly useful, for example, when planning and executing reaching and grasping movements, or movements in which a tool such as a hammer needs to contact a recipient object such as a nail. Ideomotor apraxic patients show a marked deficit in their ability to perform tasks heavily reliant on intrinsic coding (Figure 1).

Phantom Limb Phenomena

Amputees frequently experience the sensation that an amputated limb is still a part of the body. This sensation is accompanied by pain in most patients. Sensation of a phantom body part may also occur following removal of other body parts, such as after breast or eye surgery. The missing limb may feel shortened or distorted, and may seem to move, gesticulate, or twitch. A number of investigations indicate that the phantom sensations arise because of decreased input from the amputated body part to the somatosensory cortex of the brain. Areas of the somatosensory cortex that are near to the original area will take over, or "remap" the cortical region that no longer has input.

The most successful treatments of phantom limb phenomena include use of a mirror box or virtual reality to create false visual feedback about the movement of the phantom limb. When the good limb moves, participants perceive another limb moving at the location of the phantom limb. Essentially, this treatment appears to work by coupling the movement of the good limb to the tactile and proprioceptive sensations of the phantom limb, thus causing a remapping of input-output connections. Spinal cord stimulation with an electric current emanating from an electrode near the spinal cord may also be an effective treatment.

Disturbances of Body Structural Descriptions

Structural descriptions represent objects' shapes in terms of their parts, and the relationship among parts; in turn, these parts are explicitly represented as shapes in their own right. A number of lines of evidence suggest that shape primitives such as cones, cubes, and cylinders may be invariant to viewpoint (i.e., recognizable from many perspectives) and can generalize to novel objects. *Work* with patients with *autotopagnosia,* a deficit in body-part localization, suggest, that there is a specialized structural description system for the representation of human body parts.

Autotopagnosia

Patients with autotopagnosia are unable to localize body parts on themselves or others, whether cued by verbal command or instructions to model the performance of another person. Thus, for example, if an examiner points to her ear, an autotopagnosic patient may be unable to point to his own ear in turn. Although the disorder is quite rare, there are several reports of patients who show the cardinal signs of the disorder: inability to localize human body parts by pointing, but no difficulty localizing parts of animals or inanimate objects. In addition, many of these patients are able to name body parts pointed to by an examiner, and to name the functions of body parts. These data indicate that a discrete representation of the human body is separate from semantic knowledge-based representations.

Several accounts propose that autotopagnosia reflects impairment in a spatial map of the body, but only relatively recently has it become clear that this map is different from the body schema discussed in the previous section. Several interesting autotopagnosic patients, for example, have been unable to point to body parts on command, but could readily point to small objects taped to the same body parts. This indicates that the patients could perform all of the spatiomotor operations needed to localize points on the body surface and to program a reaching movement to those locations, but could still not identify discrete body parts. One of these patients had difficulty matching pictures of body parts (e.g., a nose and a different nose) when they were presented from different views. This suggests that this patient was unable to access stored representations of body parts that would permit him to judge whether the parts are the same even with occlusion of specific details. These and other data indicate that these patients suffer damage to a representation that encodes the appearance and topological location of body parts; that is, a "structural description" of the body.

Gerstmann's Syndrome

One circumscribed instance of a deficit in body structural descriptions may be seen in Gerstmann's syndrome, a relatively common group of symptoms that occurs frequently after left inferior

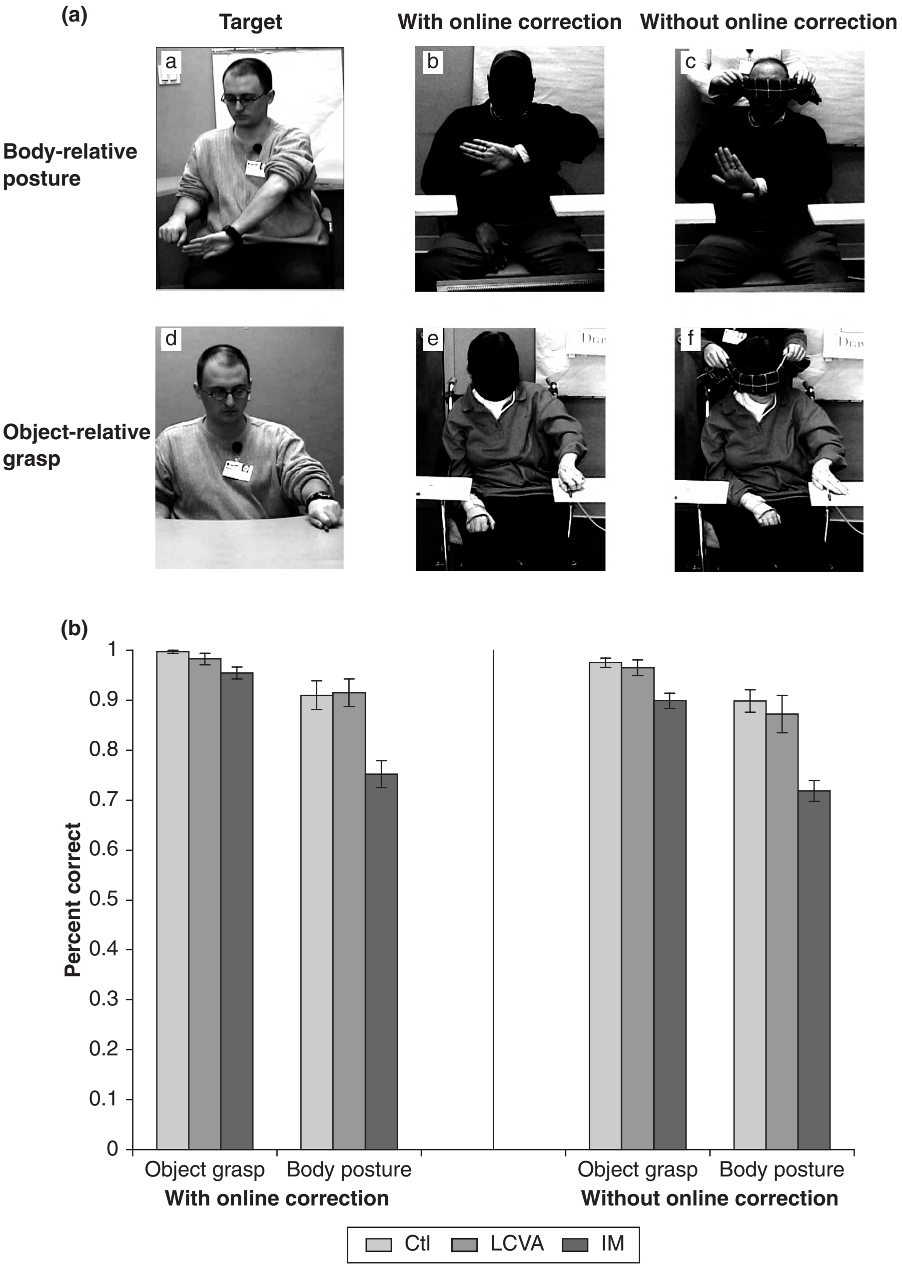


Figure 1 Ideomotor Apraxia

Source: Jax, S., Buxbaum, L. J., Moll, A. D. (2006). Deficits in movement planning and intrinsic coordinate control in ideomotor apraxia. *Journal of Cognitive Neuroscience, 18*(12), 2063–2076.

Notes: (a) Examples of performance of patients with ideomotor apraxia, from a study by Steven Jax and colleagues on a task requiring imitation of body-relative postures (top row) and object-relative grasps (bottom row). Left column: target posture. Middle column: examples of performance with full vision, where online correction is possible. Right column: performance without vision of the hand or target, where online correction is not possible. (b) Data from the task shown above. Light gray data bars are healthy control subjects (Ctl), medium gray are patients without apraxia who have suffered left-hemisphere cerebrovascular accidents (LCVA), and dark gray are patients with ideomotor apraxia (IM). Note the poor performance of apraxic patients for body-relative postures both with and without online correction, as well as for object-relative grasps without online correction. Ideomotor apraxics have difficulty positioning the body parts with respect to one another (*intrinsic coding,* explained in text), and are abnormally dependent on visual feedback.

parietal stroke. Although there is ongoing disagreement about whether Gerstmann's syndrome is a "true" syndrome, the components of the disorder are commonly observed together, and include a writing disability (dysgraphia), an inability to distinguish right from left, a lack of understanding of the rules for calculation or arithmetic (dyscalculia), and an inability to identify fingers (finger agnosia). The symptoms frequently occur in the context of a number of other neurological impairments, including *aphasia* (language disturbance), but have also been reported in pure form.

Gerstmann himself believed that a deficit in knowledge of the body was the common core underlying these symptoms. He based this hypothesis on observations that right–left disorientation is often particularly marked regarding the hands and fingers, finger differentiation is required for writing, and fingers play an important role in early learning of arithmetic operations in children, as well as in counting in some adults. Given that patients with the syndrome are impaired with others' bodies and not just their own, Gerstmann argued that knowledge of one's own body was necessary to understand another body. However, given the strong resemblance of the finger agnosia observed in Gerstmann's syndrome to the deficits exhibited with the entire body in the autotopagnosia syndrome, it remains quite possible that Gerstmann's syndrome reflects a deficit in mental manipulation of body structural descriptions.

Emerging Issues

Recent advances in research on monkey physiology indicate that there are prefrontal and parietal brain regions containing so-called mirror neurons that fire both when the monkey observes another's actions as well as when the monkey performs the same actions. Thus, these neurons may encode equivalence between the bodies of self and others. These interesting observations are likely to be the tip of the iceberg in our understanding of the encoding of body perception by the human brain. The range and variety of disorders of body perception indicate that there is considerable depth, complexity, and variability in human body processing. However, many details remain to be elucidated. It is known, for example, that some regions in the human temporal lobe—the extrastriate body area and fusiform body area—respond to the sight of human bodies, but lesions in these regions do not commonly cause problems with body perception. What, then, is the cognitive function of these regions? Conversely, as reviewed here, evidence from patient studies indicates that there are three dissociable types of body representation, each causing distinct body representation disorders, yet more work remains to clarify the precise brain regions that mediate these representations, how the representations act alone or in concert in the context of various tasks, and whether there are evolutionary antecedents as would be revealed in studies of the monkey brain.

Laurel J. Buxbaum

See also Agnosia: Tactile; Attention: Disorders; Body Perception; Cortical Reorganization Following Damage; Embodied Perception; Mind and Body; Mirror Neurons; Phantom Limb; Proprioception

Further Readings

Benke, T., Luzzatti, C., & Vallar, G. (2004). Hermann Zingerle's "Impaired perception of the own body due to organic brain disorders." 1913. An introductory comment, and an abridged translation. *Cortex, 40*(2), 265–274.

Buxbaum, L. J., & Coslett, H. B. (2001). Specialized structural descriptions for human body parts: Evidence from autotopagnosia. *Cognitive Neuropsychology, 18*(4), 363–381.

Carruthers, G. (2008). Types of body representation and the sense of embodiment. *Consciousness and Cognition, 17*(4), 1302–1316.

Committeri, G., Pitzalis, S., Galati, G., Patria, F., Pelle, G., Sabatini, U., et al. (2007). Neural bases of personal and extrapersonal neglect in humans. *Brain, 130*(2), 431–441.

Gerstmann, J. (1957). Some notes on the Gerstmann syndrome. *Neurology, 7,* 866–869.

Jax, S., Buxbaum, L. J, & Moll, A. D. (2006). Deficits in movement planning and intrinsic coordinate control in ideomotor apraxia. *Journal of Cognitive Neuroscience, 18*(12), 2063–2076.

Orfel, M. D., Robinson, R. G., Prigatano, G. P., Starkstein, S., Rüsh, N., Bria, P., et al. (2007). Anosognosia for hemiplegia after stroke is a multifaceted phenomenon: A systematic review of the literature. *Brain, 130*(12), 3075–3090.

Ramachandran, V. S., & Hirstein, W. (1998). The perception of phantom limbs. The D. O. Hebb lecture. *Brain, 121*(9), 1603–1630.

Schwoebel, J., & Coslett, H. B. (2005). Evidence for multiple, distinct representations of the human body. *Journal of Cognitive Neuroscience, 17*(4), 543–553.

BRAILLE

Braille, described in this entry, is a useful system for reading and writing by touch that was developed by Louis Braille in 1829. Before that time, some of the available systems attempted to mimic the shapes of letters. Braille is derived from a 6-millimeter (mm)-high dot matrix and does not copy letter shapes. The use of embossed letter shapes was not a successful strategy because evidence indicates that Braille patterns are far more discriminable than letter shapes, according to Jack Loomis. Although it is certainly possible to use touch to read embossed print shapes, the patterns need to be much larger than Braille characters to be useful. One would need to enlarge embossed letters to at least 1 centimeter (cm) to make them easy to feel.

Braille is an abstract code that is derived from a 2 (across) by 3 (down) rectangular matrix of embossed dots. The letters A through J are represented by the top 2 × 2 cells of this matrix, and the rest of the alphabet is produced by adding dots to the bottom two cells. Thus, the entire alphabet is produced by varying combinations of the six dots. Many blind people learn to identify the individual characters at rapid rates when they learn the code early in life.

Reading Braille Text by Hand

Most blind people are not fluent readers of Braille. The reason for this is rather simple: Namely, most people lose their sight later in life rather than in early childhood. As one ages, the frequency of visual impairment increases, and a common cause of visual impairment is diabetes, often accompanied by neuropathy and reduced tactile sensitivity. It is difficult to learn to read as an adult, and the difficulty is magnified for learning to read with one's fingers. Acquisition of skill in reading Braille is most readily accomplished in an educational environment where a child has access to good educators who have been trained in this area. However, this is less common today, given mainstreaming of visually impaired students. Former visual print readers also have considerable difficulty making the translation to reading Braille because Braille text is highly contracted. According to Susanna Millar, Braille text uses a large number of contractions to save space and time, and older visual habits may show interference effects and cause difficulty because of the presence of visual reading habits. Learning to read visually entails connecting sound patterns with visual patterns. Millar, however, contends that reading Braille requires that one attend to dot density patterns or patterns of textures, rather than letter shapes.

According to Millar, slow Braille readers tend to try to treat each character as an individual pattern or shape, and show vertical tracing patterns. For example, a letter B is represented by two vertical dots, and naive readers may move their fingers up and down as they feel the dot pair. These "scrubbing" behaviors interfere with rapid reading, which can reach more than 200 words per minute (wpm) in skilled, long-term readers. The novice is also likely to trace the shape of each character, and has a great deal of difficulty reading entire Braille words. However, skilled readers show rapid scanning of the lines of Braille, with few hesitations. Note that tactile acuity is increased in persons without sight, according to K. Sathian. Some of the difficulty that late blind persons have with learning to read Braille may be motivational. Older individuals often have difficulty coping with their loss of sight and the related restrictions in mobility. The necessity to acquire new mobility skills may limit the time available for learning to read Braille.

Laterality and Braille

The most common way that blind people read Braille is with the index finger of the right hand for reading and scanning the text. Most blind people use two hands when reading Braille skillfully. The right hand is generally used for reading the text and the left hand is used for keeping place and finding the next line of Braille as the reader completes scanning of the preceding line.

The large literature on laterality effects in vision provides evidence that the left brain excels at verbal tasks and the right hemisphere is better suited for pattern perception. This literature has led some educators of the blind to insist that students use their left index fingers for scanning Braille. This cumbersome reading method is questionable at best. It assumes that reading Braille primarily depends on the ability to perceive the individual patterns that constitute the Braille. Evidence indicates that rapid Braille reading depends on perception of texture patterns, or dot density, according to Millar. In any event, reading Braille is a complex skill that includes perceptual and higher-level cognitive, verbal components.

Multimodal Influences and Cross-Modal Plasticity

Researchers have adopted a variety of methods of studying Braille, and these have included the involvement of sighted persons and blind persons. Although the results of studies with sighted persons are interesting and informative, an important caveat should be remembered. Sighted persons are not skilled at reading Braille, and few, if any, sighted individuals can read normal Braille text by hand. For example, one finds less than 47% correct reading of two-letter uncontracted Braille words by sighted individuals according to Morton Heller, but 100% correct for two-letter words by blind persons.

Some evidence indicates that identification of Braille is enhanced in sighted individuals when they are allowed sight of the hand as it feels the Braille characters, even when they cannot see the dots themselves, according to Heller. This advantageous combination of the senses may derive from providing blindfolded sighted persons with spatial reference information about the location of the patterns in space, and the horizontal orientation. Patterns such as letters or Braille are defined by orientation, and it is easy to confuse Braille configurations when they are tilted. For example, if one rotates a Braille C (. .) by more than 90 degrees, it becomes a B (:).

Recently, researchers have been interested in how the brain processes Braille characters in blind and sighted individuals. According to Lofti Merabet and colleagues, the visual cortex is active when blind people process Braille. This indicates a reorganization of the brain owing to experience. Moreover, evidence indicates visual cortical involvement in the sighted according to Sathian and Simon Lacey. It is presently not clear exactly how the functioning of the visual cortex may differ in blind and sighted individuals, but future research will likely provide clarification of this question.

New Technology

A variety of new devices facilitate Braille output from computers, and input of Braille to the computer for blind persons. These devices are expensive, but extremely useful to the fluent reader of Braille. In addition to refreshable Braille displays that permit immediate translation of text on a computer screen to lines or pages of Braille, screen readers will read the text aloud that appears on a computer screen for the blind user. Text may be scanned using sophisticated optical character recognition (OCR) technology. Moreover, electronic Braille printers have become more widely available. However, out of convenience and cost considerations, some blind persons still use old technology (e.g., the Perkins Braille typewriter or other mechanical devices). Although voice recorders, scanners, and screen readers are helpful, they are not complete substitutes for a system of reading and writing.

Morton A. Heller

See also Low Vision; Visual Disorders: Blindness

Further Readings

Heller, M. A. (1993). Influence of visual guidance on Braille recognition: Low lighting also helps touch. *Perception & Psychophysics, 54,* 675–681.

Loomis, J. M. (1981). On the tangibility of letters and Braille. *Perception & Psychophysics, 29,* 37–46.

Merabet, L. B., Pitskel, N. B., Amedi, A., & Pascual-Leone, A. (2007). The plastic human brain in blind individuals: The cause of disability and the opportunity for rehabilitation. In J. J. Rieser, D. H. Ashmead, F. F. Ebner, & A. L. Corn (Eds.), *Blindness and brain plasticity in navigation and object perception* (pp. 23–41). New York: Lawrence Erlbaum.

Millar, S. (1997). *Reading by touch.* London: Routledge.

Sathian, K. (2000). Practice makes perfect: Sharper tactile perception in the blind. *Neurology, 54,* 2203–2204.
Sathian, K., & Lacey, S. (2007). Visual cortical involvement during tactile perception in blind and sighted individuals. In J. J. Rieser, D. H. Ashmead, F. F. Ebner, & A. L. Corn (Eds.). *Blindness and brain plasticity in navigation and object perception* (pp. 113–125). New York: Lawrence Erlbaum.

BRAIN IMAGING

One of the fundamental challenges in understanding the human central nervous system is characterizing how its various functions are localized and how they interact with one another. Imaging technologies enable neuroscientists to visualize the structure of the human brain and to measure biological activity through the skull, revealing the living brain at work. Functional localization of human perception and cognition are conventionally achieved either by recording the activity during behavior or by measuring the deficit in performance after lesioning specific regions.

At least six different families of techniques are used to investigate different aspects of human brain function: positron emission tomography (PET), functional magnetic resonance imaging (fMRI), electroencephalography (EEG), magnetoencephalography (MEG), transcranial magnetic stimulation (TMS), and lesion analysis. Each technique provides different information about brain function and affords particular strengths and weaknesses. The most spatially precise methods for localizing brain responses during mental activity are PET and fMRI. Both of these methods take advantage of the local coupling between neuronal activity and the hemodynamic response; specifically, local increases in neuronal activity lead to local changes in cerebral blood flow, blood volume, and blood oxygen concentration. Neurovascular coupling provides an indirect measure of brain activity that is fundamentally limited by the characteristics of the hemodynamic response on which the activation signal is based. Nevertheless, these methods have revolutionized the field of cognitive neuroscience during the past 10 to 20 years primarily because the activation results can be visualized as an appealing three-dimensional image that can be interpreted by the expert and layperson alike.

Finer temporal information about the millisecond time course of brain activity is achieved through EEG and MEG techniques. EEG is sensitive to scalp-surface electrical activity produced by synchronous neural firing in the brain, and MEG is sensitive to scalp magnetic fields directly generated by this electrical activity. These methods are commonly used to measure brief changes in electrical potentials or magnetic fields that are time-locked to a perceptual or cognitive event. However, the sources of the electrical generators giving rise to the transient evoked changes have to be estimated using modeling techniques, and so spatial mapping typically requires the experimenter to make certain assumptions about the underlying neural activity.

Demonstrating an association between neuronal activity and performance using any of these four methods does not necessarily mean that the brain region of interest is *required* for perception or cognition, but merely that it is engaged by the mental process. The final two methods, TMS and lesion analysis, are best suited to proof of causality (i.e., demonstrating what regions of the brain are necessary for sensation, perception, or cognition), as well as for studying the circuitry and connectivity of the brain. They both concern the suppression of neural activity, by electrical interference or brain damage, respectively. This entry describes in more detail the most popular noninvasive tools for investigating the neurophysiological basis of perception: PET and fMRI.

Positron Emission Tomography

PET is a radiotracer imaging procedure using nuclear technology to create two- and three-dimensional maps of neural activity in the brain. The term PET refers to a general method for measuring tissue metabolism by detecting the local tissue concentration of biologically important compounds, such as glucose or water, that have been labeled with small amounts of radioactivity, such as C11 or 015. A machine called a cyclotron labels these specific compounds with the positron-emitting isotopes that are then injected into the participant's bloodstream. By using different compounds, PET can quantify the amount of blood flow, oxygen and glucose metabolism, and drug concentrations in the tissues of the working brain. Sensors in the PET scanner detect pairs of gamma rays emitted indirectly by the positron-emitting radioisotope as the compound accumulates in different regions of the brain and body. These data are used to produce two- or three-dimensional images of the distribution of the drug molecules or tissue metabolism.

Blood flow and oxygen and glucose metabolism reflect the amount of brain activity in different regions and enable scientists to learn more about the physiology and neurochemistry of the working brain. Brain function is typically measured using a "fast" blood flow method. This particular technique commonly uses an isotope of oxygen incorporated into water that is injected into the blood stream and quickly perfuses the brain tissue. The concentration of the isotope builds up over a period of about 30 seconds while the participant is presented with stimuli or is performing a task. Safety guidelines on the radiation dosage typically restrict the type of participants that can be recruited, often excluding females of childbearing age, for example. Guidelines also limit experiments to a small number of scans, thus permitting the presentation of only a small number of experimental conditions. Another disadvantage of PET concerns the spatial specificity of its functional localization. Although the functionally related blood flow signal is local, it is not possible to determine how precisely it is localized to the cortical grey matter because upstream and downstream large diameter blood vessels can also display detectable changes. Spatial specificity is often reduced further during analysis by spatially smoothing the image data (up to 18 millimeters [mm]) and averaging across different brains to improve the signal-to-noise ratio. Consequently, there has been a shift during the last few years away from the use of PET for neuroscience and toward fMRI. This rapidly evolving technique conveys certain advantages including lower cost, greater accessibility, lack of radiation burden, enhanced spatial precision for localization, and the possibility of conducting single-subject investigations.

Functional Magnetic Resonance Imaging

This major branch of MRI refers to a method for generating two- and three-dimensional images of neural activity that are based on the measurement of minute changes in blood oxygenation levels in the brain. A photograph of a typical fMRI experimental setup is shown in Figure 1. Functional MRI is just one of a family of techniques developed from nuclear spectroscopy. These methods rely on the resonant properties of different atomic nuclei, and so, unlike PET, the signal contrast has an intrinsic origin. The most common signal is derived from the behavior of hydrogen nuclei (protons) because hydrogen (in the form of water, H_2O) is a major component of living tissue. In fact, the human body consists of between 45% and 75% water. Negatively charged electrons orbiting the atomic nucleus generate a spinning charge that exerts its own small magnetic field. When a participant is exposed to an external magnetic field, there is a net increase in the number of protons that are aligned with the direction of the static field.

One of the main components of an MR scanner is its huge superconducting magnet. For example, 3 Tesla refers to a field strength that is 60,000 times greater than that of the earth's magnetic field. The orientation of the net magnetization of the protons can be flipped perpendicularly by applying a short burst of radiofrequency energy that has the same resonant frequency as the hydrogen nucleus itself (i.e., 42.58 megahertz (MHz)/Tesla). When the radiofrequency pulse ceases, the protons lose energy by emitting their own radiofrequency signal that is received by a conductive field coil placed around the participant. This measurement is reconstructed to obtain a set of two-dimensional slices that can be used to create a three-dimensional image. MRI provides exquisitely detailed images not because of the high energy of the individual proton signals that are available to make the MRI signal, but because of the huge number of protons found in body tissue. The strength of the static magnetic field is important because it determines the number of excess protons and, hence, the strength of the MR signal. For example, at 1.5 Tesla, the total number of excess protons within a volume of tissue that measures 2 × 2 × 5 mm is about 6 million billion! A 3 Tesla scanner would produce an even greater excess of proton signals, hence generating brain images with an even better signal-to-noise ratio.

The rate of the return to equilibrium magnetization is determined by the magnetic state of the

Figure 1 The Whole-Body, 3 Tesla Philips Intera MR Scanner

Source: Photograph taken by D. A. Hall.

Notes: This scanner is based at the Sir Peter Mansfield Magnetic Resonance Imaging Centre at the University of Nottingham, United Kingdom. The participant is lying on the patient bed with his head positioned inside the radiofrequency receiver coil and the experimenter is moving the bed into the bore of the MR scanner. In this auditory fMRI experiment, the participant was required to discriminate between different sound stimuli that are presented over headphones. Verbal responses are not possible during fMRI and so responses were recorded using a 10-button keypad that is resting on the participant's leg.

surrounding tissue. fMRI capitalizes on the fact that deoxygenated hemoglobin affects this magnetic environment, with the rate of decay becoming more rapid as the neighboring concentration of deoxyhemoglobin increases. Given that the metabolic demands of neuronal activity determine the oxygenation state of the blood in localized regions of the brain, the MR signal provides a convenient measure of neuronal activity. The component of the MR signal that is associated with the underlying neural activity is referred to as blood-oxygenation-level-dependent (BOLD) contrast. BOLD contrast provides the basis for fMRI. The size of the BOLD signal is small, accounting for only about 1 to 2 % of the overall signal variance, and so experiments involve averaging image data acquired during several tens of repeated presentations of the stimulus. Physiological noise, such as the cardiac and respiratory signals and movement of the cerebrospinal fluid, are the main contributors to the MR signal variance.

In reality, the relationship between neuronal activity and the BOLD signal is complicated because the contrast is based on combined changes in relative blood flow, volume, and oxygenation. The precise coupling between the BOLD response and neural activity is still in debate, but it is generally accepted that the BOLD signal is reasonably well correlated with population synaptic activity (including inhibitory and excitatory activity). Not all of the BOLD signal comes from hemodynamic changes within the smaller venules and capillaries that are most closely located to the site of neuronal activity. Some of the BOLD signal comes from larger draining veins and so the BOLD activation maps probably extend a few millimeters beyond the underlying neural activity. The spatial resolution of the functional image itself is also in the order of several millimeters, with fMRI scans generally having a spatial resolution of about 3 × 3 × 3 mm.

Brain Imaging Studies

The functional data that are acquired during PET and fMRI are too noisy to construct an activation map from a single brain scan and so experiments typically involve repeating the same stimulus or task many times while the participant undergoes a series of brain scans. Figure 2 (see also color insert, Figure 37) illustrates some of the key aspects of a brain imaging experiment using an example taken from auditory fMRI. The upper time line represents one of the simplest paradigms for mapping neural activity, one in which the experimenter seeks to contrast the brain response during a stimulation condition (such as listening to a particular sound stimulus) with that during a baseline condition (such as listening to no sound stimulus at all, or to a different sound stimulus). A single PET scan takes many seconds to produce and so each scan tends to capture the ongoing neural activity pattern that is sustained across a stimulus condition of about 30 to 90 seconds in duration. In contrast, a single fMRI scan can be acquired rapidly (within 1–2 seconds), and so a series of scans are typically acquired over time. fMRI is therefore more sensitive to both transient and sustained hemodynamic changes and provides much more flexibility about the choice of stimulus duration (from a brief event lasting 100 milliseconds (ms) to a longer epoch of 30 seconds or so). Another difference between the two brain imaging methods concerns the total number of scans that can be acquired for each participant. PET is limited by safety guidelines that limit the overall radioactivity dosage for an individual and a maximum of 12 scans is typical for most experiments. In contrast, the BOLD signal is generated via the endogenous contrast mechanism provided by proton resonance, and there are no known safety risks. The main practical constraint for fMRI scanning time is the total amount of time that the participant is able to lie still and concentrate on the task (between 10 and 30 minutes).

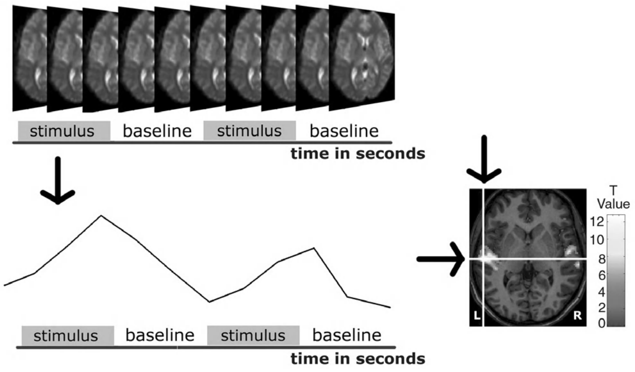


Figure 2 Several Key Aspects of an (Auditory) fMRI Experimental Design and Image Analysis

Source: Original figure created by D. A. Hall.

Notes: The upper time line represents a segment of the experiment showing how repeated scans are acquired at regular intervals over time, while the participant listens to an alternating sequence of sound stimulation and baseline conditions. Each scan is represented by one horizontal slice through the brain. In this example, each stimulus condition lasts about 9 seconds, with scans occurring every 3 seconds. The lower time line presents a schematic diagram of the stimulus-correlated variability in the MR signal measured in arbitrary units. Such a time course is representative of a voxel (individual point) that responds strongly to the sound stimulus. The spatial distribution of the activated voxels is shown in the right panel, with the most significant voxel highlighted by the crosshairs. The color scale reflects the statistical significance of the activation across the auditory cortex. Note that the activation is overlaid onto the anatomical scan, which has a finer spatial detail than the functional scans shown in the upper time line. In this figure, the activation reflects the greater contralateral auditory cortical response to a sound stimulus that was presented to the right ear. See also color insert, figure 37.

Activation maps of the BOLD signal are generated from a statistical analysis of the image data. However, statistical analysis follows a series of image processing steps that are carried out after the data have been acquired by the MR scanner. Each image processing step is rather computationally intensive and must be completed off-line. The analysis makes the assumption that the position of the brain is constant throughout the time series, so first, the effects of small head movements over time are corrected so that each and every scan represents the same set of brain slices. If the pattern of neural activity is to be directly compared across different participants, or if the data from different participants are to contribute to a group analysis, then the shape of the individual brains must be spatially transformed to match a common stereotactic space. An official standard brain template is used to provide an international reference that includes both a set of three-dimensional coordinates and anatomical labels for the corresponding brain regions. Most often, image analysis uses a parametric statistical approach. The General Linear Model (GLM) is applied to each point in the image (individual points are called voxels). The model describes the changes in the amplitude of the MR signal over time as a linear sum of the different experimental effects (corresponding to the BOLD signal), the expected confounding effects, and the residual variability. The pattern of neural activity is estimated by contrasting the GLM parameter estimates at each voxel for the different stimulus conditions (for example, by applying a T-test to the stimulation and the baseline conditions). In the example shown in Figure 2 (also color insert, Figure 37), the "baseline" scans are subtracted from the "stimulation" scans using a T-test procedure and the regions of the brain that show a statistically significant difference in signal between the two conditions ($p < 0.05$) is defined as functionally active. The statistical significance of the contrast is calculated using classical statistical inference. Often the color scale of the activation map represents the statistical value (such as the T value) for each voxel that exceeds the chosen significance threshold (see Figure 2; see also color insert, Figure 37). The activation maps obtained using fMRI represent a relative increase in the response to one stimulation condition compared with another. The maps do not represent an absolute measure of neural activity. In contrast, activation maps obtained using PET can represent an accurate estimate of the actual change in local regional cerebral blood flow. Activation maps are often overlaid onto a high-resolution anatomical brain scan that is acquired using MRI. The anatomical scan maximizes the contrast between gray and white matter in the brain by choosing a different set of gradient and radio frequency pulse parameters of the MR scanner compared with those used in BOLD imaging. The anatomical scan typically has a spatial resolution of at least 1 × 1 × 1 mm and is used to facilitate the interpretation of precisely where in the brain the activation is located.

Deborah Hall

See also Cortical Organization; Evoked Potential: Audition; Evoked Potential: Vision; Neural Representation/Coding; Neuropsychology of Perception; Oscillatory Synchrony; Transcranial Magnetic Stimulation

Further Readings

Frackowiak, R. S. J., Friston, K. J., Frith, C. D., Dolan, R. J., Price, C. J., Zeki, S., et al. (Eds.). (2004). *Human brain function* (2nd ed.). San Diego, CA: Academic Press.

Jezzard, P., Matthews, P. M., & Smith, S. M. (Eds.). (2001). *Functional magnetic resonance imaging: An introduction to methods*. Oxford, UK: Oxford University Press.

C

CAMOUFLAGE

The most common kinds of camouflage make one thing appear to be two, two things one, and so on. The occurrence of camouflage predates human history, in the sense that abundant examples exist of concealment and deception in nature. Some plants are often mistaken for stones, some birds have feathers that look like bark, and some insects resemble twigs (called *mimicry*). There are also natural textures that closely match their surroundings (called *blending* or *crypsis*), or forms that are so visually broken that they defy recognition (*disruption*) (see Figure 1; see also color insert, Figure 13). This entry describes natural camouflage, military applications, Gestalt psychology and camouflage, and more recent developments.

Natural Camouflage

Scientists became particularly interested in camouflage in nature in the 19th century because they saw it as tangible evidence of the theory of natural selection, which Charles Darwin had proposed in 1859. Research of the subject advanced in the 1890s, when the U.S. artist and naturalist Abbott H. Thayer claimed to have determined why so many animals are colored lightest on their undersides, where they are the least sunlit, and darkest on their upper parts. It is now widely believed that animals with such *countershading* (also known as Thayer's law) appear to be less solid and less dimensional because their markings cancel out the shading effects of the overhead sun.

During the Spanish-American War, Thayer proposed unsuccessfully that countershading should be applied to U.S. ships. In 1902, he patented the idea, which he claimed to have derived from the coloration of a seagull. His research intensified and, in 1909, he and his son (Gerald H. Thayer) published a groundbreaking volume, titled *Concealing Coloration in the Animal Kingdom,* in which (among other things) they suggested ways to use "protective coloration" for military purposes. In addition to countershading, the book described and illustrated natural and humanmade examples of blending, disruption, and mimicry.

Military Applications

During World War I, partly through Thayer's revelations, hundreds of artists, architects, theater designers, and others were deployed by various armies as consultants in wartime deception. The French, who were the first in history to establish a military unit for that purpose, referred to this practice as *camouflage,* and a camouflage advisor was a *camoufleur*.

Initially, most ground-based or field camouflage was concerned with concealing artillery from airborne observers, who relayed the location of targets to their own long-distance artillery. In 1917, however, British attention shifted to the heavy loss of their ships from attacks by German submarines (called U-boats). This led to a new way of thinking about naval camouflage as distinct from field

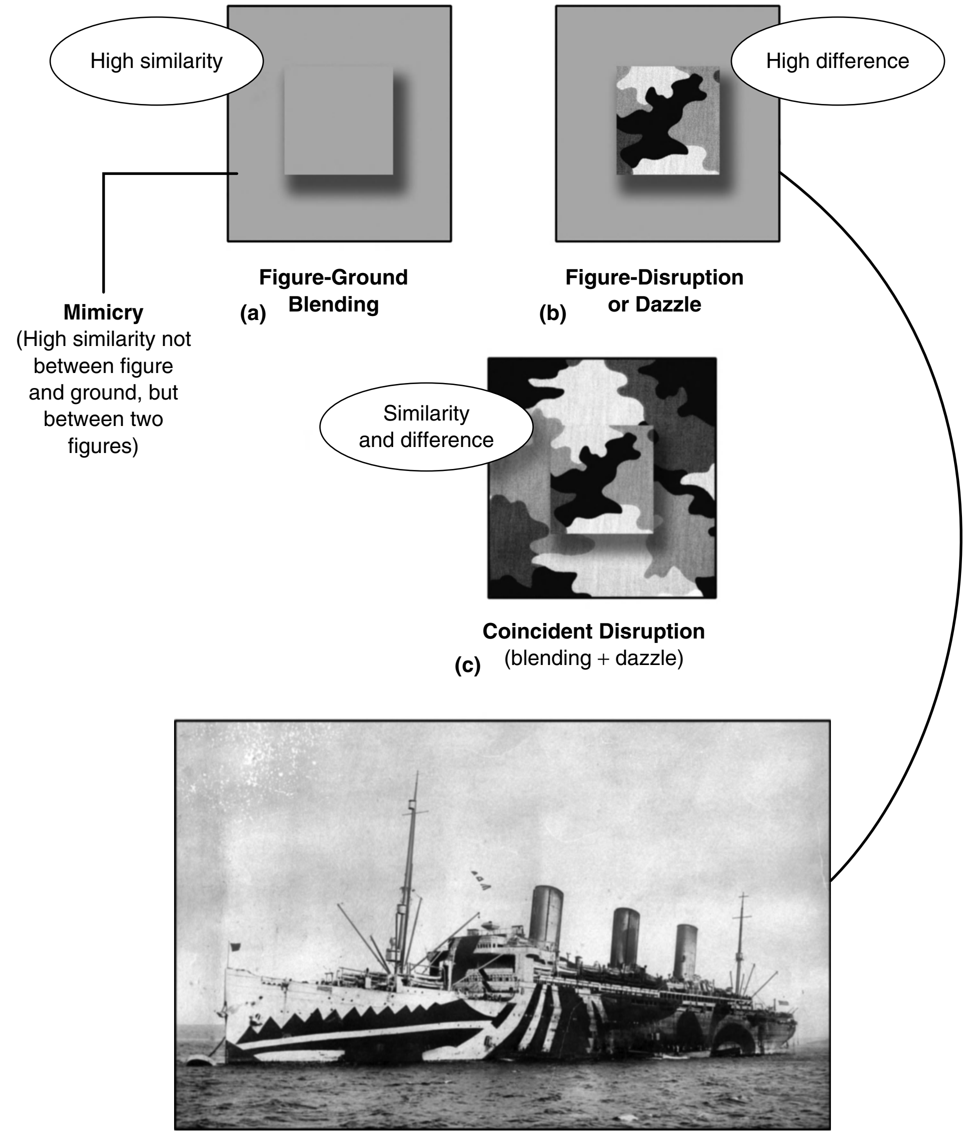


Figure 1 Three Categories of Camouflage

Source: Illustration © by Roy R. Behrens, 2008. Reprinted with permission.

Notes: Although there are any number of ways to categorize examples of camouflage, perhaps the three most common kinds are figure-ground blending (a), in which an object blends in with its setting (also sometimes known as crypsis); figure disruption ([b] and photo below it), in which the surface of the figure is broken up by highly contrasting components, a method that was used widely for ship camouflage during World Wars I and II; and coincident disruption (c), in which the figure is broken up, but parts of it blend in with the background. In the system shown here, mimicry is said to be a subcategory of blending or high similarity camouflage. See also color insert, Figure 13.

camouflage, with the result that the British began to apply erratic, abstract patterns to the surfaces of their ships. First proposed by the British artist Norman Wilkinson (who called it *dazzle painting*), the purpose of this method was not to conceal the ship, but rather to throw off the accuracy of the U-boat gunner, who was positioned at a considerable distance and had to aim ahead of the target. It was never scientifically confirmed to what extent this actually worked, but the U.S. government was sufficiently convinced that it deployed its own camouflage expert (the artist Everett Warner) and adopted a parallel strategy in the last year of World War I.

Gestalt Psychology and Camouflage

The findings by Thayer and others about camouflage and its military applications were more or less concurrent with the birth of a branch of psychology called Gestalt theory. Founded in Germany around 1910 by three psychologists (Max Wertheimer, Kurt Koffka, and Wolfgang Köhler), this movement was chiefly concerned with the psychology of vision and with problem solving. Notably, the Gestaltists investigated *figure-ground phenomena* (the extent to which a shape is seen, in relation to various backgrounds), the conditions of which are of central concern in the visibility of objects. Gestaltists identified the *principles of perceptual organization* (similarity and proximity grouping, good continuity, closure, etc.), which did not differ greatly from the principles Thayer had earlier called the *laws of disguise* (e.g., countershading, figure-ground blending, and shape disruption). Gestaltists also brought attention to the holistic or synergistic aspects of experience, as expressed by the oft-quoted maxim (derived from Aristotle), "the whole is greater than the sum of its parts."

English translations of writings about Gestalt psychology became available in the mid- to late-1920s. In some of their publications, the Gestaltists used examples of camouflage as proof of their earlier findings. They were particularly interested in children's puzzles called *embedded figures* (also known as a *camouflaged figures* or *puzzle pictures*) in which one is asked to find a figure or other shape in a purposely complex setting, a task that one of their students, Kurt Gottschaldt, used extensively in his research of problem solving. Later, in 1940, the British zoologist Hugh B. Cott (who had been a camoufleur during World War I, and a camouflage instructor in World War II) published a comprehensive book on *Adaptive Coloration in Nature,* in which he argued persuasively that embedded figures are ubiquitous in natural camouflage. He referred to this as *coincident disruptive coloration,* by which he meant the combined use of blending and disruption.

More Recent Developments

Throughout World War II, those trained in art-related fields continued to be called on to contribute to camouflage, both military and civilian. By the end of the war, this had lessened, partly because of developments in surveillance technologies that were not based on vision (such as radar). In addition, because the effectiveness of dazzle camouflage had never been proven, it was no longer cost efficient to painstakingly design and apply bewildering abstract patterns to thousands of ships.

After World War II, military applications of camouflage evolved, on the one hand, toward electronic, nonvisual deception (such as the illusive technology of the Stealth bomber). On the other hand, initially as an adaptation to the jungle setting of the Vietnam War, the camouflage patterns on military uniforms became increasingly visual, to the point that some might regard them as advertising rather than as concealment.

In recent years, there has been a heightened interest in camouflage, as shown by the adoption of camouflage patterns in civilian clothing styles. There has also been increased discussion about the pertinence of camouflage (both literal and metaphorical) to art, architecture, philosophy, and literature. Additionally, in the biological sciences, there has been a resurgence of research about natural camouflage.

Roy R. Behrens

See also Gestalt Approach; Magic and Perception; Object Perception; Perceptual Organization: Vision; Perceptual Segregation; Pictorial Depiction and Perception

Further Readings

Behrens, R. R. (2002). *False colors: Art, design and modern camouflage.* Dysart, IA: Bobolink Books.

Behrens, R. R. (2009). *Camoupedia: A compendium of research on art and camouflage.* Dysart, IA: Bobolink Books.

Blechman, H. (Ed.). (2004). *DPM: Disruptive pattern material: An encyclopedia of camouflage: nature, military, culture.* London: DPM.

Cott, H. B. (1940). *Adaptive coloration in animals.* London: Methuen.

Gregory, R. L., & Gombrich, E. H. (Eds.). (1973). *Illusion in nature and art.* New York: Scribner's.

Newark, T. (2007). *Camouflage.* New York: Thames & Hudson.

Thayer, G. H., & Thayer, A. H. (1909). *Concealing coloration in the animal kingdom.* New York: Macmillan.

CAUSALITY

One is tempted to believe that perception is limited to the reconstruction of meaningful objects such as rocks or chairs on the basis of raw sensory information. The topic of perceptual causality shows that one should resist that temptation and be open to widen the scope of perception. Subjective causality is by no means a mere reflection of physical causality, but is our interpretation of whether one event is causing another event. It often approximates reality and sometimes deviates substantially from it. The function of this subjective and immediate sense of causality is to make sense of the world and to guide our actions. Albert Michotte has in many ways become the father of perceptual causality by arguing that the phenomenal impression of causation is a fundamental building block of perception and that causality is perceived in a direct manner that is not qualitatively different from that of perceiving rocks or chairs. Before describing the work of Michotte and his followers, this entry looks more closely at causality and its context of event perception.

Efficient Versus Final Causes

Causality can be conceived in two different manners, one relating to internal events, such as goals, opinions, and cultural settings. The other relates to the observation of external events that happen in the world. Aristotle proposed this distinction between causation from within and external causation. He called the former *final causation* and the latter *efficient causation.* Take, for instance, the event of a player's leg making contact with a ball, upon which the ball proceeds to move into the direction of a box. In terms of efficient causation, one could say that the leg has caused the ball to move. The leg moved first and somehow imparted its motion or impetus to the ball. We observed one motion as the cause of another subsequent motion.

David Hume has extensively discussed this type of causality. Besides temporal succession, one has to assume additional mechanisms that distinguish a true cause from coincidental succession. For instance, in the case of a collision of two billiard balls, the ricochet motion of the target ball has been caused by the cue ball via such intermediate mechanisms as friction and energy transfer. A soccer fan would not agree to this description. She would describe the same event as a forward player intending to score a goal. The intention causes the player to aim for the goal, and this intention causes the ball to move. Unlike the efficient cause, the final cause relates to an involved observer, for whom the simple chain of events (1) intention (2) leg movement (3) ball movement is no longer a chain. The intention does not end when the leg starts to move, nor does it disappear when the leg has stopped moving.

Cause and effect cannot easily be narrowed down in temporal or spatial terms. Although the conditions that must be met to perceive an efficient cause can be described with comparative ease, the conditions for perceiving a final cause are more intricate because they do not fit into straightforward descriptive categories. Not surprisingly, psychologists have put their primary focus on the efficient cause although the final cause is in many ways more important for us as acting individuals. In the world of a human actor, actions are caused by the intended environmental effects. The final cause for our fingers flipping the light switch is our intention to turn on the lamp. This intention stretches to the point in time where the lamp is on, and thus the succession of cause and effect is no longer valid. To make things even more difficult, the technological advances of our time require us to perceive causal relations between goals and events that appear to be dissimilar. Where our ancestors pulled a mechanical lever to set a switch, we achieve the same effect by means of a mouse click on the computer or even a verbal input.

Phenomenal Reality of Efficient Causes

Since the days of Hume, our understanding of an efficient cause has been shaped by the idea of associationism. We see A as the cause of B because

A precedes B with regularity. Turning the switch is perceived to cause the light to go on because the switch action has become associated with the light. In contrast to this associative concept of causality, Michotte put forth a fundamentally different concept of visual causality, claiming that we have direct phenomenal access to causality. It requires neither learning nor conscious or unconscious inferences. He was able to demonstrate that observers perceive the motion of two simple objects as causally related as long as surprisingly few boundary conditions are met. Take, for instance, the case of an actual billiard shot. The cue ball moves toward the stationary target ball and abruptly stops at the moment of collision, instantly setting the target ball into motion. Michotte discovered that impoverished renditions of such an event suffice to convey the impression of causality. The motion of a painted dot toward another painted dot on a screen and the subsequent motion of the second dot are readily and immediately perceived as causally connected.

Michotte did most of his experiments using a turntable. Two lines were drawn on a rotating disk, which was covered by a mask except for a small slit that served as a viewing port. Changing the distance between the circular lines at different stages would produce different dot motions. Whenever the first dot moved toward the second dot, followed by motion of the second dot, the stimulus was perceived as "dot 1" launching "dot 2." Despite the artificial nature of the stimuli generated with this method, a stable impression of causality was readily produced. Depending whether the first dot stopped after contact with the second, Michotte distinguished among a launching effect proper, entrainment, launching across a spatial or a temporal gap, and others. Some of these effects are illustrated in Figure 1.

Infants readily perceive launching and entraining, which suggests the innate nature of phenomenal causality. Michotte's insights remain valid, albeit with considerable elaborations. Efficient causality has turned out to be more complex than thought by Michotte. In some cases, observers are reluctant to make causal attributions where they should according to Michotte, and in other cases, observers perceive causality even outside the conditions he put forth for launching and entrainment. For instance, judgments of causality change depending on the perceptual attitude and experience of the observer. Observers who adopt an analytic attitude tend to

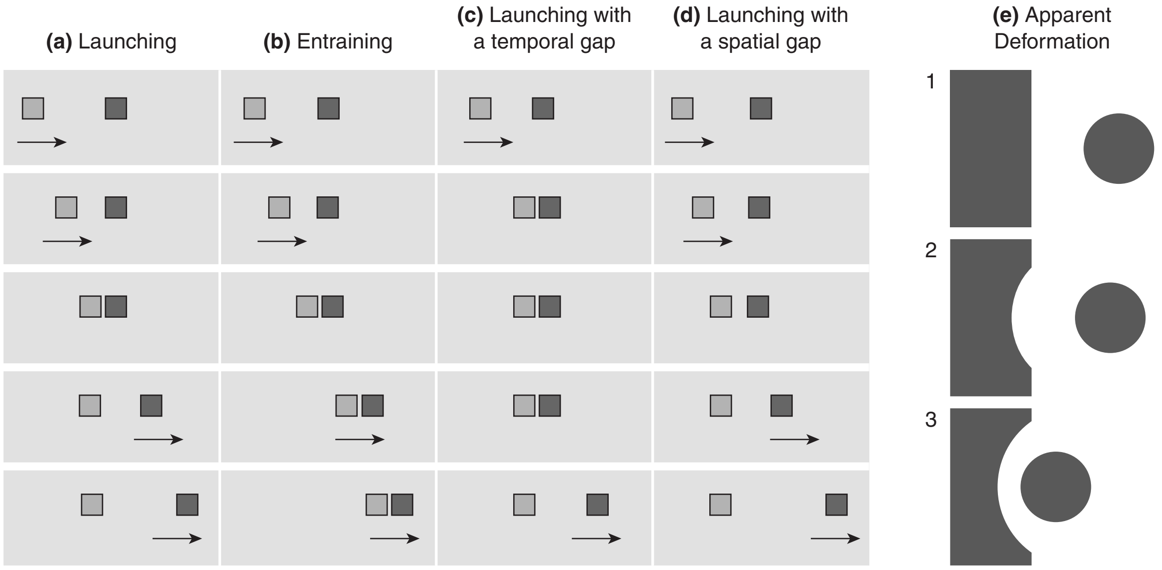


Figure 1 Schematic Illustration of Michotte's Stimuli

Notes: (a) Launching, (b) entraining, (c) launching with a temporal gap, (d) launching with a spatial gap. In (e), the observer has the impression that the motion of the circle causes the deformation of the rectangle although neither launching nor entrainment is the case.

describe the standard launching stimulus as a sequence of independent movements, rather than cause and effect. Conversely, the sequence of events shown in column (e) of Figure 1 is readily seen by most observers as the gray circle putting a dent into the grey rectangle although the two never touch. This suggests that the boundary conditions for perceptual causality may be even wider than originally thought.

Intuitive Physics

From an evolutionary point of view, one could argue that true causal relations have only been discovered by the perceptual system inasmuch as they have been internalized to secure the survival of the species. Whenever erroneously perceived causal relations did not have any adverse consequences, errors were not corrected. Thus, our direct access to efficient causes should be limited by the needs of the hunter and gatherer who moved around in a world determined by the laws of classical mechanics. According to the field of intuitive physics, the approximation of perceptual causality to physical causality has not progressed any further than what was known about physics in ancient Greece. For instance, when asked to indicate the trajectory of an object dropped from a moving carrier, many observers believe that it will fall straight down. Accordingly, when being moved on a conveyor belt, many adults and most children will release a ball directly above the stationary target on the ground when intending to hit it. They overshoot the target because they act as if they were stationary and the ball moved straight down. Instead, it moves down and in the direction of self-motion. After a few trials, the behavior adjusts quickly, but the causal relations that are expressed by Newton's laws remain opaque. The understanding of classical mechanics seems indeed limited to pre-Newtonian theory. The direct access to causation has clear limitations that are well described by intuitive physics. In other words, human observers have internalized causal relations that hold in the environment, but only to a rough approximation. This approximate internalization would also explain the inclination to perceive causal structure in mere graphic illustrations.

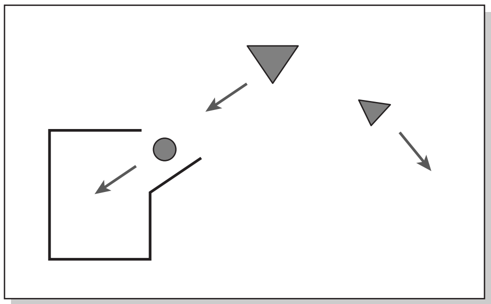


Figure 2 Geometric Figures Used by Fritz Heider and Georg Simmel in a Cartoon

Notes: The arrows indicate the direction of motion and were not present in the experimental stimuli. Movements of the figures are seen as actions and the figures take on particular traits.

Overestimation of Action Goals

When it comes to final causes, a new domain of perception opens up, one that is sometimes dramatically removed from physical causality. The laws of physics and mechanical causality are being bent dramatically when it comes to the perception of causes for willed events. For example, in the case of ballistic movements, a large number of observers believe that a cannon ball accelerates after it has left the cannon. Similarly, observers judge a thrown ball that increases its speed after leaving the pitcher's hand to look more natural than a ball that decelerates in accord with physical law. This surprising but stable finding can be explained by the projection of an intention into the object itself. Rather than limiting the effect of the intentional cause to the movement of the effector—the arm that throws the ball—the effect migrates to the ball. In Aristotle's terms, a person's goal to throw the ball to a particular target causes the ball to move, not just while it is being held in his or her hand but also beyond this point on its way to the target.

Final causality also leaves the domain of perception proper and enters the realm of emotion and cognition. We spontaneously "see" intentions, emotions, and social relationships in simple movements of geometrical objects as Fritz Heider and Georg Simmel have demonstrated. When the geometric drawings in Figure 2 are animated, as indicated by the arrows, observers spontaneously describe the actions in social terms. They may readily see the large triangle as a villain chasing the circle while the small triangle is the friend of the circle (apparently a female), and the large rectangle is a house. We cannot help but experience inanimate motion within a framework of social meaning.

Interestingly, even stationary objects tend to be interpreted in a causal manner that considers the history of the object's deformations, as Michael Leyton has pointed out. A crumpled piece of paper readily reveals that its prior state was uncrumpled and that it has been put into its crumpled state by an actor.

Heiko H. Hecht and Dirk Kerzel

See also Event Perception; Intentionality and Perception; Phenomenology (Philosophy); Social Perception

Further Readings

Hecht, H. (2001). Regularities of the physical world and the absence of their internalization. *Behavioral and Brain Sciences, 24,* 608–617.

Heider, F. (1944). Social perception and phenomenal causality. *Psychological Review, 51*(6), 358–374.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57,* 243–259.

Leyton, M. (1992). *Symmetry, causality, mind.* Cambridge: MIT Press.

Michotte, A. (1963). *The perception of causality.* London: Methuen.

Schlottmann, A., Allen, D., Linderoth, C., & Hesketh, S. (2002). Perceptual causality in children. *Child Development, 73*(6), 1656–1677.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Science, 4*(8), 299–309.

Wagemans, J., van Lier, R., & Scholl, B. J. (2006). Introduction to Michotte's heritage in perception and cognition research, *Acta Psychologica, 123*(1–2), 1–19.

White, P. A. (2006). The causal asymmetry. *Psychological Review, 113*(1), 132–147.

CELL PHONES AND DRIVER DISTRACTION

Operating a motor vehicle is an activity engaged in by more than one billion people on an annual basis. Given that each year there are more than one million fatalities worldwide associated with motor vehicle accidents, driving represents one of the more risky activities undertaken on a regular basis by adults between the ages of 25 and 65. Estimates suggest that as many as 80% of accidents on the roadway are the result of some type of driver distraction, and it has been noted that one of the basic road errors is a failure to see and react to another road user in time. Driver distraction associated with visual processing can be divided into two general categories: Situations where drivers fail to look at roadway hazards in the driving environment (i.e., failures to look or fixate upon an object) and situations where drivers look at objects in the driving scene, but show degraded processing of the driving environment because attention is directed elsewhere (i.e., failures to see or attend to the object). This entry describes the relationship between cell phones and driver distraction.

Driver distraction associated with a failure of visual processing can be attributed to situations where drivers fail to direct their eyes to hazards in the roadway. This source of impairment can be the result of distractions from outside the vehicle (e.g., an electronic billboard with sudden onsets or movement that results in a reflexive orientating of the eyes to the billboard) or to distractions within the vehicle (e.g., dialing or text messaging on a cellular phone). In both situations, driving is impaired because the eyes are diverted from the roadway and the information necessary for the safe operation of the vehicle is not fixated on for a sufficient duration to allow for its processing. This results in a source of driver distraction wherein the driver fails to look at critical information in the driving environment. Research indicates that the concurrent performance of ancillary visual tasks often compromises driver safety; however, some drivers are more impaired than others. Compared with more experienced drivers, novice drivers are less able to interleave the performance of a concurrent visual task with driving. For example, novices tend to allocate longer epochs of visual processing to a secondary task than do more experienced drivers who exhibit frequent and short glances between the two tasks. In sum, the first source of distraction occurs if the driver does not look at an object, and thus, he or she cannot see and react in a timely fashion.

Driver distraction can also be caused by attention being diverted from driving to some other activity (e.g., a cell phone conversation), resulting in a form of inattention blindness. In such cases, critical information in the driving environment may be fixated upon, but the driver does not "see" the information because attention is directed elsewhere. For example, when drivers converse on either a hand-held or hands-free cell phone, they fail to subsequently recognize as much as 50% of the information that they would have recognized had they not been conversing on a cell phone. In addition, the impaired recognition is equivalent for both items of high and low relevance to safe driving, suggesting that there is little or no semantic prioritization of the information that is impaired by the cell phone conversation (i.e., drivers are as blind to a billboard on the side of the road as they are to a pedestrian standing on the curb). Moreover, recordings of brain activity indicate that cell phone conversations suppress the initial registration of information in the driving scene. In particular, the amplitude of the P300 component of the event-related brain potential elicited by imperative events in the driving environment (e.g., brake lights) is reduced by half when drivers are conversing on a cell phone. Thus, drivers using a cell phone fail to see information in the driving scene because they do not encode it as well as they do when they are not distracted by the cell-phone conversation. In addition, the diversion of attention from the driving environment increases the time it takes for drivers to respond to changes in the driving scene (e.g., lead vehicle braking and red-onset traffic lights). In situations where the driver is required to react with alacrity, the data indicate that drivers using a cell phone will be less able to do so because of the diversion of attention from driving to the phone conversation. The net result is that activities such as conversing on a cell phone increase the likelihood and severity of accidents on the roadway.

Several other multitasking activities share surface similarities with conversing on a cell phone but do not lead to significant driver distraction. For example, listening to radio broadcasts or books on tape does not impair driving. Evidence also indicates that speech production results in significantly greater interference than does speech comprehension. These findings are important because they demonstrate that listening to verbal material, by itself, is not sufficient to produce the dual-task interference associated with using a cell phone while driving. Moreover, in-vehicle conversations do not interfere with driving as much as cell phone conversations because there is a greater ability to synchronize an in-vehicle conversation with the processing demands of driving than there is with a cell-phone conversation.

Whereas the fact that driver distraction is associated with failures to fixate on objects in the driving environment and failures to attend to objects that are fixated, the mechanisms underlying the latter source of interference are not well understood. In particular, models of dual-task performance that invoke concepts such as "attentional resources" or "code conflict" would seem to predict that an auditory/verbal/vocal task such as conversing on a phone should not interfere with a visual/spatial/manual task such as driving. It is possible that the interference stems from a central processing bottleneck associated with the concurrent tasks of driving and talking on a cell phone. Further specifying the mechanisms for this source of dual-task interference is an ongoing focus of research.

David L. Strayer and Joel M. Cooper

See also Attention: Divided; Attention: Effect of Breakdown; Change Detection; Eye Movements: Behavioral

Further Readings

Levy, J., Pashler, H., & Boer, E. (2006). Central interference in driving: Is there any stopping the psychological refractory period? *Psychological Sciences, 17,* 228–235.

McCarley, J. S., Vais, M., Pringle, H., Kramer, A. F., Irwin, D. E., & Strayer, D. L. (2004). Conversation disrupts scanning and change detection in complex visual scenes. *Human Factors, 46,* 424–436.

Strayer, D. L., & Drews, F. A. (2007). Cell-phone induced inattention blindness. *Current Directions in Psychological Science, 16,* 128–131.

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied, 9,* 23–52.

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular phone. *Psychological Science, 12,* 462–466.

CHANGE DETECTION

Change detection is the noticing of change in the world around us. For example, when driving in traffic, the vehicles around us constantly change their positions, and sometimes their speed and direction as well. To avoid collisions, we must notice any such changes and respond to them. More generally, the ability to detect change is important for coping with everyday life; humans (and most animals) are consequently quite good at it. However, our knowledge of how changes are detected remains incomplete. The progress that has recently been made owes its existence to two related developments: (1) a realization of some of the confusions and assumptions built into our previous ideas about change, and (2) new methodologies that allow it to be more effectively isolated and studied.

An important step is to clarify the meaning of the terms themselves. As used here, *change detection* is restricted to the noticing of a change (i.e., the observer seeing that a change *exists*) via the use of vision. This can include the related abilities of identifying the change (i.e., seeing *what* it is), as well as localizing it (i.e., seeing *where* it is), although these abilities likely involve somewhat different mechanisms.

An adequate understanding of change detection has been difficult to achieve. Part of this is because of the nature of change itself. Although the concept of change appears simple, attempts to formalize it have shown otherwise. For example, change requires that some aspect of an object remain constant but another aspect does not, a situation that has not been completely resolved by present-day philosophers. Furthermore, our intuitions about change detection are often highly inaccurate. For example, we generally believe that we could easily detect any change in front of us provided that its size is sufficiently large. But we can be amazingly "blind" to such changes, failing to detect them even when they are large, repeatedly made, and are expected. Such *change blindness* is a phenomenon strikingly at odds with our intuitions about how change detection should work. However, such counterintuitive results have taught us much about what change detection is and how it works, as this entry describes.

Background

Insight into change detection has been obtained from a variety of sources. Some of the earliest discoveries occurred in Hollywood studios in the 1930s, where filmmakers found that an audience could fail to notice a sudden change (e.g., the appearance of a new character) if this occurred at the moment they made an eye movement (or saccade) from one side of the screen to the other, or when they blinked their eyes in response to a sudden noise. Filmmakers also discovered their own blindness to changes made during a film cut (e.g., the change in length of cigarettes in different shots), with such continuity errors often going unnoticed until long after the film had been released.

Controlled investigations of this phenomenon began in the 1950s. One line of research examined simple displays composed of about a dozen dots with random positions; observers were found to be surprisingly poor at determining if one of these dots had changed in a later presentation of that image. This work ultimately formed the basis for the proposal of a limited-capacity visual short-term memory (VSTM), an important part of current theories of visual perception.

A later wave of studies in the 1990s extended this in several important ways: images were more realistic, repeating changes were often used, and change blindness was also induced in other ways, such as making changes during an eyeblink. These developments allowed change detection to be more deeply investigated, and connections made to other aspects of visual perception. One of the more important findings in this regard was the discovery that *attention is needed to see change.*

Methodologies

All studies of change detection use essentially the same design: An observer is first shown an initial stimulus (e.g., a picture of a harbor with some ships in it), followed by a modified version of this (e.g., the same picture with one of the ships removed). The ability of the observer to detect the change is then measured, usually in terms of speed or accuracy.

In real life, the motion signals produced by a change automatically attract attention, causing it to be easily seen. (For example, waving at your friends usually makes them see you right away.) To study the mechanisms involved in change detection, the process is slowed by eliminating this automatic drawing of attention. This can be done in several ways.

Gap-Contingent Techniques

In this approach, the initial image is presented for a short time (e.g., a half second), followed by a brief blank field or mask (typically about 100 milliseconds [ms]); the modified image is then shown (see color insert, Figure 14). The change therefore occurs during the gap between the original and modified stimuli. If this gap is longer than about 80 ms, changes can be extremely difficult to see—indeed, some can require more than a minute to be seen, even though continually repeated.

The explanation for this is that the “flicker” created by the gap overwhelms the local motion signals created by the stimulus change, making it impossible for them to automatically draw attention to its location. The observer must then carry out a time-consuming scan of the image until the change is found.

Saccade-Contingent Techniques

Here, the initial image is again presented to the observer, but the transition to the modified image is made during an eye movement (or *saccade*) to a new location. The stimulus on a retina is smeared during a saccade. Thus, from the point of view of the retina, the pattern encountered here is this: initial image—smeared image—modified image. This is similar to that encountered in the gap-contingent techniques, and so induces similar amounts of change blindness.

Shift-Contingent Techniques

In this approach, the initial image is shown to the observer for a short time, and the modified image then shown immediately afterward, with its location shifted somewhat—for example, moved right or left a few degrees of visual angle. Changes made under these conditions are again difficult to detect, even when they are large and expected.

The reason for change blindness here is much like that for the gap-contingent techniques: the shifted image means that there is mismatch at most points in the original and modified image, generating motion signals that swamp those created by the change itself. Note that this is closely related to the blindness found for changes made during film cuts, where the signals created by the mismatch of the images swamp the local signals created by the change, requiring a time-consuming attentional scan of the image.

Gradual-Change Techniques

These techniques have the initial image gradually fade away, with its place taken by the modified image. If the transition occurs over several seconds, the change will be extremely difficult to see. The reason for change blindness here is essentially the same as that for the other conditions: the ability of the motion detectors to guide attention to the location of the change has been eliminated, in this case, by making the change slowly enough that the motion detectors don’t fire.

Implications for Perceptual Mechanisms

The failure to see large changes under various experimental conditions can be explained by the proposal that attention is needed to see change. However, the story does not end here. New distinctions and new questions arise, many of which have more general implications for how we see. These include the following.

Change Versus Motion

Work on change detection motivates a more careful distinction between change and motion. Here, *motion* is the temporal variation in a property at a given point in space. For example, the flow of muddy water in a stream can be described by its velocity at each point. There is no object here, no sophisticated structure—only unstructured “stuff” extending over part of space. Simple motion detectors in the human visual system respond to exactly such variation.

In contrast, *change* refers to the transformation of a particular object or event over time—for example, the wing movement of a bird overhead. Here, there is a constant structure (the bird) in which a transformation (wing movement) occurs.

Attention is important here, being used to form a coherent representation that persists during its transformation.

In this view, any dynamic variation in the world is picked up by two separate perceptual systems: one for *motion* (variation in a particular location) and one for *change* (variation in a particular structure). Both systems will generally be in play. One of the key insights of recent work is the need to separate the contributions of each system, so that their operation can be studied independently.

Mechanisms of Change Detection

Another set of issues involves the mechanisms of change detection itself, and how attention enters into their operation. An important finding is that only 3 to 4 items can be seen to change at a time. Indeed, it appears that the items held by attention may be not independent, but may instead be parts of a single structure that corresponds to a single object.

Recent work is beginning to explore the components involved. Contrary to our intuitions, change detection need not be an elementary process. It can be divided—at least conceptually—into a sequence of distinct operations: (a) load the information into a memory store (such as VSTM), (b) put it into a coherent form, (c) hold it across a temporal gap, (d) compare the contents of memory to the new stimulus, (e) unload this memory, and (f) shift processing to the next candidate item(s). Results indicate that distinct mechanisms for at least some of these do exist.

Scene Perception

Given that our representation of a scene contains no more than a few coherent structures at a time, an important question is why we have the impression that we see everything in coherent form. (Note that much of a scene is represented—at least while our eyes are open—but this is not necessarily in coherent form.)

Recent work supports the view that scene perception is based on a dynamic process, rather than on an accumulation of detailed information. More precisely, attention appears to be allocated on a "just-in-time" basis, with the representation of an object formed at the moment it is needed, and dissolved once attention has been withdrawn. Longer-term visual memories are still needed to coordinate this process. But these may contain relatively sparse amounts of information—perhaps just the information from the items currently attended, along with some properties from the locations used in the management process.

Implications for Individual Differences

Change detection can help explore the general mechanisms involved in visual perception, as well as how particular individuals (and cultures) perceive the world around them. One approach to this is based on the *priority* given to items in a display. Given that attention is first allocated to interesting items, the speed at which a change is detected can indicate if the associated item is considered interesting. For instance, observers addicted to a substance such as alcohol or cannabis are faster and more likely to detect changes to items associated with that substance than are observers who are not addicted.

Individuals also differ in the *way* they encode items, something that depends at least partly on training. Experts at U.S. football are better able to spot meaningful changes to football scenes than are novices, indicating that the experts have learned to see aspects of the scene not perceived by others. In addition, some studies suggest that U.S. residents may be less able to detect changes in the surrounding context (e.g., the surrounding elements in a scene) than are Japanese, indicating a difference in the encoding of foreground and background elements.

Finally, individuals can also differ in the *mechanisms* at their disposal. For example, some individuals can have a "gut feeling" that something is changing, even though they do not yet have a "picture" of the change. Although there is still dispute about the basis of this phenomenon, the fact that observers can show differences in the way they experience change may provide an interesting new perspective on how we perceive our world.

Ronald A. Rensink

See also Attention: Effect on Perception; Attention: Selective; Eye Movements: Behavioral; Motion Perception; Object Perception; Visual Memory

Further Readings

Brockmole, J. R., & Henderson, J. M. (2005). Object appearance, disappearance, and attention prioritization in real-world scenes. *Psychonomic Bulletin & Review, 12,* 1061–1067.

Dmytryk, E. (1984). *On film editing.* Boston: Focal Press.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology, 23,* 420–456.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–280.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology, 53,* 245–277.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review, 5,* 644–649.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9,* 16–20.

Werner, S., & Thies, B. (2000). Is "change blindness" attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition, 7,* 163–173.

COCHLEAR IMPLANTS: CONTROVERSY

Cochlear implants have been dogged by controversy since the earliest stages of their development. In the mid-1970s, when French otologist Claude-Henri Chouard, a pioneer in implant surgery, was implanting on average one deaf patient per month, the French Deaf community protested against the procedure, staging rallies in 1977 and 1978. About that same time, another implant pioneer, Graeme Clark of Australia, found himself greeted by Australian Deaf protesters who resented the negative portrayal of deaf people and their lives that Clark promoted in the media as he sought to sell the public on his research. He reported that he was "surprised" to find that deaf people were upset with his efforts to "cure" them. This entry describes the controversy regarding cochlear implants.

Many in the Deaf community—meaning those people who participate in Deaf culture, have a deep appreciation of Deaf history and traditions, and use a signed language to communicate—do not consider deafness a medical condition in need of a cure. They believe deafness is a necessary precursor to Deafness, and Deafness is about being a member of a cultural and linguistic minority group. This smaller Deaf community, it should be noted, is distinct from the millions of people who have some degree of hearing loss. For instance, although some 13 million U.S. residents have varying degrees of hearing loss, only about 600,000 of them call themselves Deaf.

These Deaf Americans, like their international counterparts, reacted negatively to cochlear implants. The Greater Los Angeles Association of the Deaf (GLAD) published a position paper on implants in 1985. That report expressed concern that the medical profession generally failed to acknowledge the existence of a Deaf perspective about deafness, let alone about implants. The National Association of the Deaf (NAD) issued a position paper on implants in 1991, condemning the use of implants in children and objecting to the oralism that the technology promotes.

It is not simply the audiological intervention, the imperative to hear, that draws the ire of many in the Deaf community. Members also object to the pressure to speak. The cochlear implant technology promotes speech, and only speech, as the expected communication method for implantees. For example, Web sites of both implant manufacturers and surgical centers stress that the goal of the procedure is to promote speech and listening. Some also recommend the complete abandonment of any signed language with a deaf child, arguing the continued use of signs will interfere with the acquisition of speech.

Many in the Deaf community view this as an attack on their language and their culture, as much as on their physicality. This is not the first time such issues have been raised. Oralists at the beginning of the 20th century believed that sign language should be banned from deaf education altogether, and only speech and lipreading should be used with deaf schoolchildren. From the point of view of many in the Deaf community, cochlear implants are a high-tech weapon in an old battle against American Sign Language, a prominent marker of Deaf culture. To many, this technology is trying to do what oral education techniques alone could not, that is, make deaf people over in a hearing image.

Culturally Deaf people have historically been quite uninterested in the project of being made over as hearing people. Rather, they took pride in their Deafness. To many, such implant technology would eliminate deafness but suppress signed language, and thereby undo Deaf culture. However, this threat to the Deaf culture was largely lost on the wider hearing public in the same decade. Many hearing people believed, and arguably still do, that deafness is a medical condition in need of a cure.

As the rates of implantation increased in the 1990s, reactions to cochlear implant technology became more strident. However, at the beginning of the 21st century, the Deaf community shifted its focus from contesting the issue with the hearing community to making implanted deaf people feel welcome in the Deaf world. That is, if implants continued to grow in popularity, the Deaf community wanted those implanted to join them, and claim their place in the Deaf world. Accordingly, the NAD revised its position paper on implants in 2000. It spoke of recognizing the "diversity within the deaf community" and stressed that the Deaf world "welcomes all individuals regardless of race, religion, ethnic background, socioeconomic status, cultural orientation, mode of communication, preferred language use, hearing status, educational background, and use of technologies." The statement also stresses that "implanted children are still deaf" and they should be encouraged to develop both speech and sign skills accordingly. It is too soon to tell whether this call to inclusion will bring implant users into the Deaf community.

R. A. R. Edwards

See also American Sign Language; Cochlear Implants: Technology; Prostheses: Visual

Further Readings

Baynton, D. (1996). *Forbidden signs: American culture and the campaign against sign language.* Chicago: University of Chicago Press.

Edwards, R. A. R. (2005). Sound and fury; or, much ado about nothing? Cochlear implants in historical perspective. *Journal of American History, 92,* 892–920.

Lane, H. (1993). Cochlear implants: Their cultural and historical meaning. In J. V. Van Cleve (Ed.), *Deaf history unveiled: Interpretations from the new scholarship* (pp. 272–291). Washington, DC: Gallaudet University Press.

National Association of the Deaf. (2000). *Cochlear implants. NAD Position Statements.* Retrieved from http://www.nad.org/ciposition

COCHLEAR IMPLANTS: TECHNOLOGY

A *cochlear implant* (CI) is an auditory prosthesis developed to treat severe-to-profound sensorineural hearing loss (SNHL). Hearing loss is sensorineural (rather than conductive) when the hearing dysfunction is located at the inner ear (cochlea) or along the neural pathway that connects the cochlea to the brain. SNHL can occur congenitally from genetic causes or from complications during the prenatal or early postnatal periods. It can also be acquired as a result of disease, injury, tumors, ototoxic drugs, noise exposure, or aging. SNHL usually involves damage to the cochlea's hair cells, which vibrate in response to sound. These vibrations induce nerve cells to send signals to the brain that cause the sensation of hearing. Damage to the hair cells is permanent and cannot be repaired by any medical or surgical means available today. When hair cell damage is relatively mild, hearing aids can compensate for the reduced sensitivity by amplifying sound. However, if the hair cells are damaged beyond responsiveness, hearing aids are ineffective and cochlear implantation becomes a viable alternative.

Rather than amplifying sound, CIs bypass the hair cells and directly stimulate auditory neural fibers connected to the cochlea with electrical impulses. The auditory neural fibers (as well as the corresponding hair cells) are organized tonotopically in the cochlea; neural fibers at the front (apical) end of the cochlea respond to the highest frequencies in the environment, and nerve cells at the back (basal) end of the cochlea respond to the lowest frequencies. CIs divide sound into frequency bands and then stimulate the cochlea tono-topically according to the intensity of the auditory input in each frequency band. Today, approximately 120,000 people worldwide have received a CI, among which more than 6,000 have two. This entry discusses how CIs work, the process of cochlear implantation, CI effectiveness, and new directions.

How Cochlear Implants Work

Components

A cochlear implant system has external and internal components. The external components shown in Figure 1, sit behind the ear and on the side of the head. A microphone, worn behind the person's ear (A), sends signals to a sound processor (not shown), a small unit about the size of an MP3 player that can be worn on a pocket or specially designed pouch. The sound processor converts incoming sounds into electrical pulses. The electrically processed sounds are sent to a transmitter (B), which is held in place by a magnet just behind the ear. These signals are transmitted from the transmitter through the skin to a receiver, implanted under the skin (C), and from there, the signals are sent to the electrodes (as many as 22) in the electrode array (D). The electrical pulses directly stimulate auditory neural fibers that help the cochlear implant recipient hear.

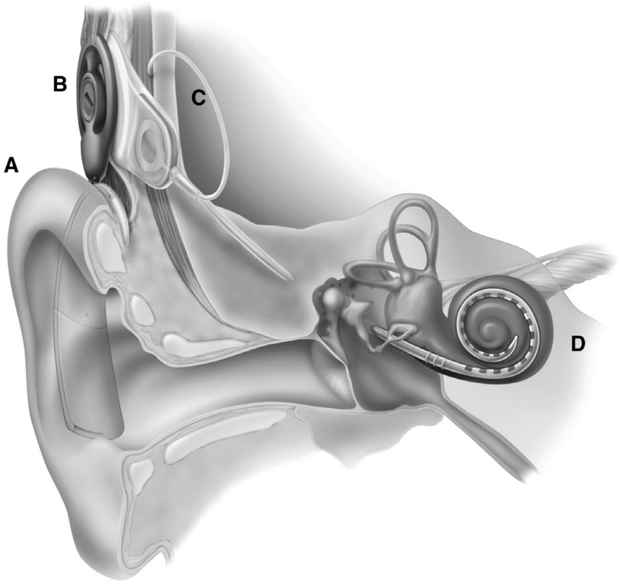


Figure 1 The External and Internal Components of a Cochlear Implant

Source: Courtesy of Cochlear Ltd.

Notes: The external components sit behind the ear and on the side of the head. A microphone, worn behind the person's ear (A), sends signals to a sound processor (not shown), a small unit about the size of an MP3 player that can be worn on a pocket or specially designed pouch. The sound processor converts incoming sounds into electrical pulses. The electrically processed sounds are sent to a transmitter (B), which is held in place by a magnet just behind the ear. These signals are transmitted from the transmitter through the skin to a receiver, implanted under the skin (C), and from there, the signals are sent to the electrodes (as many as 22) in the electrode array (D). The electrical pulses directly stimulate auditory neural fibers that help the cochlear implant recipient hear.

Sound Processing

Although different types of sound processing strategies exist across the manufacturers of cochlear implants, the basic principles for each strategy are similar. Specifically, the environmental sounds are sent through a bank of filters (i.e., electronic circuits that remove specific frequency components of the signal) based on their frequency or pitch. Low frequency sounds or components of sounds are maintained in the low-pass band filters, and high-frequency sounds are represented in the high-pass band filters. Each filter then outputs a series of electrical pulses or pulse trains, which are electrical representations of the input signal or sounds. These electrical pulse trains are used to stimulate the auditory neural fibers.

Cochlear Implantation

Candidacy Criteria

To receive a CI, adults and children over the age of 2 must have a severe-to-profound hearing loss (i.e., ability to minimally detect the loudness of signals that are comparable to the loudness of a lawn mower or a subway train), and children 12 to 24 months old must have a profound hearing loss. Also, both children and adults must demonstrate through audiological testing that their spoken word comprehension is poor when using hearing aids. Individuals must have an intact auditory nerve and no medical contraindications for surgery. Finally, it is recommended that children undergo a psychological evaluation to identify any confounding factors that may affect the successful use of a CI. In the United States, the FDA has approved cochlear implantation for children 12 months of age or older. There is no upper age limit for implantation.

Surgery

Skin incisions are made to provide access to the mastoid process (protruding bone behind the ear) while preserving the blood supply of the skin. The implant electronics and induction coil are placed in a pocket under the skin. A hole is drilled in the mastoid so the electrode array can be inserted into the middle ear through the facial recess and then into the cochlea through an opening made anterior and inferior to the round window membrane of the cochlea. The insertion is performed to position the electrode into the scala tympani of the cochlea. Following the insertion, the hole into the cochlea is sealed with connective tissue and the skin incision is sutured.

Programming the Cochlear Implant

Approximately one month after surgery, the CI is programmed for each individual and his or her hearing capabilities. Typically, the threshold levels (i.e., the lowest signal intensity a person detects) and the comfort levels (i.e., strongest signal a person can tolerate comfortably) are determined for each electrode along the electrode array. After programming, the electrodes stimulate the auditory nerve at near threshold levels for soft sounds, and stimulation will be greater with louder sounds but never greater than comfort levels. Within the first 6 months, the audiologist modifies these levels regularly as the patient adapts to the new signal.

For children and adults who can communicate when they are hearing the electrical signals, programming the device can be accomplished efficiently and reliably. For infants, however, programming the implant is more challenging because infants cannot reliably report when they are hearing. To overcome these difficulties, implant manufacturers have developed their devices and software such that neural activity from the auditory nerve can be directly measured in response to electrical stimulation. This neural activity can provide an indication of loudness levels where the infant first hears the input signal. Audiologists can then use this information to program the CI.

Cochlear Implant Effectiveness

Variability in Outcomes

Most adults who receive CIs had normal hearing during childhood. They can use linguistic knowledge to help interpret the new sound input from the CI. As a result, most adults with CIs have near-normal scores on a variety of outcome measures after a period of adaptation. Many adult CI users report that at first voices sound more high-pitched and mechanical than normal. This improves with adaptation; however, CIs never restore hearing to the natural quality that adults remember. Congenitally deaf children, on the other hand, do not have a memory of spoken language. Thus, there is a large amount of variability in outcomes among children with CIs: Some children do very well whereas others receive little benefit from the CI.

Predictors of Outcome Variability

There is an enormous amount of variability in language outcomes among children with CIs. Some of this variability can be predicted by several demographic factors. One of the most important is the age at which children receive a CI. Earlier implantation predicts better language development. For example, recent investigations suggest that toddlers who receive an implant before 1 year of age learn words more quickly than do toddlers who received their implant at 1 to 2 years of age. Electrophysiological measures, which can measure the development of natural auditory responses from infancy to adulthood, have revealed that there is likely an important window of time during early childhood in which it is crucial to acquire spoken language.

Another factor that predicts language outcomes is the amount of hearing at the time of implantation. Children with more hearing before implantation tend to acquire language better than do children with less hearing before implantation. Characteristics of the CI and the surgery also matter. Outcomes are affected by the number of electrodes inserted and the depth of CI insertion.

After surgery, the type and amount of speech therapy is crucial. Children who receive frequent speech therapy support develop better speech and spoken language skills than do those who receive little or no speech therapy. Children who undergo therapy that focuses more exclusively on spoken language tend to perform better on tests of speech perception and speech production than do children enrolled in therapy that combines sign and spoken language. Also, children's language skills improve as they gain more experience with the CI.

In addition to demographic predictors of outcome, researchers have recently investigated other factors that might contribute to success. For example, children who have better cognitive, motor, and lipreading skills before receiving a CI have been shown to have better speech and spoken language outcomes.

Effectiveness for Other Sound Perception

Recently, researchers have begun to investigate the effectiveness of CIs for perceiving other types of sound. Studies of music perception and production have shown that CI users are good at rhythm but not melody. Research on different qualities of speech has shown that CI users have a difficult time identifying individual talkers and dialects. Pitch cues in speech, such as rising pitch at the end of a question or lexical tone in some Asian languages, as well as complex non-speech sounds in the natural environment, such as animal or insect sounds, are also challenging for CI users.

New Directions

New Configurations

Until recently, all patients received a CI in only one ear and the other ear was saved for future technology. Now, some investigators are reporting advantages in locating sounds and speech understanding in noise with bilateral implantation, and many people believe that the benefits of binaural hearing may outweigh the benefits of waiting for a new technology. As a result, increasingly more patients are requesting a second CI, and some receive two CIs simultaneously.

CIs eliminate all acoustic ("natural") hearing in the implanted ear. For CI candidates who have some residual hearing for low-pitched sounds, it is desirable to preserve their acoustic hearing while using the CI to gain access to high-pitched sounds. One way to accomplish this is to implant one ear and fit the other ear with a hearing aid. Recently, CI manufactures have created a "short electrode array" that preserves the hair cells in the cochlea that respond to low-pitched sounds, allowing a mix of electrical and acoustic hearing in a single ear. Some evidence indicates that both of these approaches result in better speech understanding than does unilateral implantation.

New Measures of CI Effectiveness

Today, measures of CI effectiveness primarily involve tests of speech and language comprehension and production during the first several years after cochlear implantation. One direction for future research is to assess the long-term effect of CIs on education level and general quality of life. Another future direction is to investigate how early deafness and subsequent cochlear implantation affect cognitive, social, and emotional outcomes.

Assessing Effectiveness in Infants

Because earlier implantation predicts more successful outcomes across several measures, the candidacy criteria for CIs have expanded to include increasingly younger children. Currently, the FDA approves implantation down to 1 year, and some surgeons are implanting infants as young as 5 months.

Cochlear implantation during infancy presents new challenges. Current clinical instruments for assessing language skills in children involve instructing the children to perform tasks that are impossible for infants, such as pointing at a named picture or repeating what the clinician says. To assess early linguistic skills in infants, clinicians must currently rely on parental reports, which are indirect and highly dependent on the parents' ability to objectively assess their child's performance.

To develop more direct measures of performance, researchers are currently adapting behavioral measures that have been used to investigate normal-hearing infants' speech perception and language skills for use in the population of infants with hearing loss. For example, one methodology assesses infants' ability to distinguish different syllables or words using the following procedure. Infants are presented with an image of a checkerboard pattern on a TV monitor and a repeating speech sound until the infant looks away from the checkerboard pattern. Their looking times are monitored by an observer and recorded on a computer.

The presentation of the checkerboard pattern and speech sound is repeated until the infant's looking time to the checkerboard pattern has decreased by 50%. Then, the same checkerboard pattern is presented but now with a different speech sound. Infants who can distinguish the new speech sound from the old one will tend to increase their looking time to the checkerboard pattern in response to hearing the new speech sound. However, infants' performance on these types of tasks is highly variable and affected by things such as fussiness and attention. Current research continues to hone these techniques to improve their reliability.

Derek M. Houston, Marcia J. Hay-McCutcheon, Tonya R. Bergeson, and Richard T. Miyamoto

See also Audiology; Audition: Disorders; Auditory Processing: Peripheral; Cochlear Implants: Controversy; Hearing Aids

Further Readings

American Speech-Language-Hearing Association. (2004). *Technical report: Cochlear implants* (ASHA Supplement 24, pp. 1–35). Rockville, MD: ASHA.

Houston, D. M., Pisoni, D. B., Kirk, K. I., Ying, E., & Miyamoto, R. T. (2003). Speech perception skills of deaf infants following cochlear implantation: A first report. *International Journal of Pediatric Otorhinolaryngology, 67,* 479–495.

Litovsky, R. Y., Johnstone, P. M., & Godar, S. P. (2006). Benefits of bilateral cochlear implants and/or hearing aids in children. *International Journal of Audiology, 45*(Supplement 1), s78–s91.

Loizou, P. C. (1998). Mimicking the human ear. *Signal Processing Magazine, IEEE, 15,* 101–130.

McDermott, H. J. (2004). Music perception with cochlear implants: A review. *Trends in Amplification, 8,* 49–82.

Miyamoto, R. T., & Kirk, K. I. (2003). Cochlear implants. In J. J. Ballenger & J. B. Snow (Eds.), *Ballenger's otorhinolaryngology head and neck surgery* (16th ed., pp. 476–488). Hamilton, Ontario: BC Decker.

Nicholas, J. G., & Geers, A. E. (2007). Will they catch up? The role of age at cochlear implantation in the spoken language development of children with severe to profound hearing loss. *Journal of Speech, Language and Hearing Research, 50,* 1048–1062.

Pisoni, D. B., Conway, C. M., Kronenberger, W. G., Horn, D. L., Karpicke, J., & Henning, S. C. (2008). Efficacy and effectiveness of cochlear implants in deaf children. In M. Marschark & P. C. Hauser (Eds.), *Deaf cognition: Foundations and outcomes* (pp. 52–101). New York: Oxford University Press.

Sharma, A., Dorman, M. F., & Kral, A. (2004). The influence of a sensitive period on central auditory development in children with unilateral and bilateral cochlear implants. *Hearing Research, 203,* 134–143.

COLOR: GENETICS OF

Normal trichromatic color vision in humans is afforded by the presence of three types of light sensitive cone photopigments in the retina. The underlying genetics of the short-, middle-, and long-wavelength sensitive (abbreviated S, M, and L) cone photopigments has been the topic of research for nearly 300 years. However, only during the past 20 years, through the use of modern molecular genetic techniques, has precise information about the genes underlying normal and defective color vision become available. This entry examines the current understanding of the genetic basis for human color vision, genetic variability in the system, and its behavioral consequences.

Genetic Basis of Human Color Vision

Cone photopigments in the human retina have two components—a protein component (opsin) and a chromophore (11-*cis*-retinal, a vitamin A derivative). Each of the cone photopigments has different spectral sensitivities; that is, they vary in the efficiency with which they absorb the different wavelengths of light. These cone photopigments differ only in their opsin component, and a different gene encodes for each of the opsin molecules. The gene encoding the S opsin is located on chromosome 7. The genes encoding L and M opsins are located on the X chromosome, in a head-to-tail tandem array (i.e., they lay right next to each other). This arrangement is the primary reason there is variation in these genes in humans. Just adjacent to this array is an important genetic regulatory element called the locus control region (LCR). The LCR is required for any of the genes in the L/M array to be expressed. The L and M opsin genes are about 96% identical, though they show only about 43% identity with the S opsin gene; many of these genetic differences are responsible for the different spectral properties of the photopigments. Mutations in these opsin genes result in differences in the structure of the associated opsin. This alters the functional properties of the photopigment; for example, the sensitivity of the photopigment to different wavelengths of light might be shifted, or the photopigment could be altogether nonfunctional. The functional consequences of genetic variation are discussed later.

The location of the L and M genes on the X chromosome means that color vision defects associated with mutations in these genes are inherited in an X-linked recessive pattern. These defects are commonly called "red-green" color vision defects. Females have two X chromosomes, whereas males have only one; therefore, in males, one mutated copy of the X chromosome is sufficient to cause a red-green deficiency, whereas in females the mutation must be present in both copies of the X chromosome. Thus, males are about 25 times more likely to have red-green defects than are females; however, nearly 15% of women will be carriers of a red-green color vision defect. Interestingly, there is rarely any behavioral indication of a female's carrier status. Thus, these red-green defects might pass through many generations without any family member exhibiting a defect.

In contrast, the autosomal (not sex-linked) location of the gene encoding the S opsin means that color vision defects associated with mutations in the S-opsin gene (referred to as "blue-yellow" color vision defects) are inherited in an autosomal dominant fashion. As a result of their autosomal location, humans normally have two copies of each of these genes; yet, a defect in a single S pigment gene is sufficient to cause a blue-yellow defect—there are no carriers. Thus, blue-yellow defects are equally prevalent in males and females, though they are much more rare than the red-green defects in general.

Genetic Variation Within Normal Color Vision

Most mammals have only a single L or M gene on the X chromosome; their color vision, therefore, is based on an S photopigment and an L or M photopigment, giving them dichromatic color vision. Most Old World primates (including humans) possess both L *and* M genes; their color vision is based on three photopigments (S, M, and L), rendering them trichromatic.

Among humans with trichromatic color vision, the number of genes in the L/M array varies from two (the minimum required to confer trichromacy) to as many as nine, with a mode of three. Curiously, other trichromatic Old World primates show nowhere near this degree of variability in gene number, even though they too are trichromatic. The relevance of the variability in gene number for human color vision is not clear, especially given evidence suggesting that only the first two genes in the L/M array can be expressed.

The location of the L/M gene array and their high homology allow for frequent unequal homologous crossovers (intermixing of the genes). This can occur between the genes (intergenic crossover) or within the genes (intragenic crossover) between two parental X chromosomes. Such recombination events are responsible for the observed variation in L/M gene number on the X chromosome. In addition, genes with "intermixed" L and M sequences are produced from intragenic crossover events. These genes encode photopigments that have spectral sensitivities intermediate of the normal L- and normal M-cone photopigments (and are sometimes called hybrid pigments). To date, there has been no measurable real-world consequence for having these hybrid pigments in individuals with normal color vision, although their presence can be revealed through careful color vision testing where subtle differences in spectral sensitivity can be revealed. The combined variability in gene number and gene sequence means that in a group of 100 males, the probability that any 2 will have identical L/M arrays is less than 2%.

Genetic Variation and Red-Green Defects

Two main genetic causes of inherited red-green color vision deficiency have been identified. The most common cause is rearrangement of the L/M genes resulting either in the deletion of all but one of the genes, or in the production of a gene array in which the first two genes both encode a pigment of the same type (L & L or M & M). The second general cause is the introduction of a mutation in either the first or second gene in the array, rendering one pigment nonfunctional. The most prevalent mutation results in the disruption of a highly conserved disulfide bond that is essential for the proper folding of the protein into its functional three-dimensional form. The same mutation, when it is present in the rod photopigment, causes an aggressive form of blindness called autosomal dominant retinitis pigmentosa.

Although all red-green deficient individuals have compromised color discrimination compared with normal trichromats, significant variation correlates with the underlying genetic variation. On one end of the spectrum are *dichromats* (individuals with a single spectral type of L or M gene). They have the worst color discrimination, as evident with a Rayleigh color match where they accept all mixtures of the red and green primaries as matches to the monochromatic yellow. The remaining individuals (who have either multiple types of L pigment and no M pigment or multiple types of M pigment and no L pigment—referred to as *anomalous trichromats*) fall along on a continuum of deficiency. In some cases, the deficiency is so mild that the individual is unaware of it only until genetic or behavioral testing; however, the most severe individuals behave almost like dichromats. This variation in behavior can be reliably linked to the underlying genetics of anomalous trichromacy, though exceptions can arise from differences in experience or performance issues on behavioral tests. The L and M photopigments differ in their peak spectral sensitivity by almost 30 nanometers (nm). Consider an anomalous trichromat with two L genes. If the L genes encode for L photopigments that are nearly identical to each other in their spectral sensitivity, the defect will be quite severe (nearly dichromatic). However, if these genes encode for an L photopigment and a hybrid photopigment (with disparate spectral sensitivity), the resultant defect can be mild. Nevertheless, it is now possible to examine the L and M genes of an individual and reliably determine whether he or she has a red-green defect and, if so, to make a prediction about the severity of the defect.

Joseph Carroll and Diane M. Tait

See also Animal Color Vision; Color Deficiency; Color Perception: Physiological

Further Readings

Gegenfurtner, K. R., & Sharpe, L. T. (1999). *Color vision: From genes to perception.* New York: Cambridge University Press.

Jacobs, G. H. (1993). The distribution and nature of colour vision among the mammals. *Biological Reviews, 68,* 413–471.

Nathans, J., Piantanida, T. P., Eddy, R. L., Shows, T. B., & Hogness, D. S. (1986). Molecular genetics of inherited variation in human color vision. *Science, 232,* 203–210.

Neitz, J., & Jacobs, G. H. (1990). Polymorphism in normal human color vision and its mechanism. *Vision Research, 30,* 620–636.

COLOR: PHILOSOPHICAL ISSUES

Why does the apple on the table before you look red? A commonsense answer: You are looking at a red apple in good natural light and you have normal color vision. It may seem hard to quarrel with this statement, but it is not particularly informative. A scientific account might go something like this: The apple selectively reflects more of the longer wavelength rays in daylight than those of shorter wavelengths. As a result, the kinds of light sensors in the retina of your eyes that are more sensitive to longer wavelengths respond more strongly than do those that are more sensitive to shorter wavelengths. This imbalance in response is coded and re-coded, exciting appropriate neural networks in the brain. Properly fleshed out, as it presumably will be in the future, such an account is indeed informative, but it would seem to lack something important. Nowhere in it is there a reference to the red quality of the apple or to your experience of that quality. The philosopher's task is to try to put the commonsense and the scientific accounts together to get a picture that accounts for the red quality and relates it to an informative causal analysis. The main philosophical attempts to do so are briefly discussed in this entry.

Physicalism, Dispositionalism, and Primitivism

Our only access to the colors of things is through our visual experience. As the *physicalist* sees it, the red of the apple is that physical property of the apple that causes us to experience red. This would seem to be its selective reflectance of light rays across the spectrum. However, the three sensor types of the retina are limited in their ability to resolve the spectral details of the light arriving at the eye. Consequently, things with different spectral reflectances can look red. A comparable situation applies to the spectra of light-emitting objects such as LEDs. The causes of experiences of red are so physically diverse that it is not clear what or whether they have anything in common.

To secure a property that all red things have in common, *dispositionalists* propose that we take the redness of an object to consist in its disposition to look red to color-normal observers under normal conditions, regardless of the physical mechanisms that underlie that disposition. Unfortunately, this proposal has problems of its own. The color appearance of some objects varies significantly with variation in the spectral composition of the light source, even if the source is considered "normal." For example, two surfaces that match in color in sunlight may look distinctly different in north daylight. In addition, the visual background and the distribution of colors in the scene will always affect the color appearance of the object to some extent, and sometimes quite dramatically. Furthermore, color-normal observers do not see colors of things in exactly the same way. So what viewing conditions and which observers shall we use to determine the "true" colors of an object? There seem to be no convincing reasons for choosing one set of conditions or observers rather than another. Once more, we are without a unitary objective feature that we could identify with redness.

Primitivists are dissatisfied with the candidates for physical redness proposed by physicalists and dispositionalists, but still want to retain the intuition that colors are properties of the material world outside of our consciousness. They propose that colors be understood as primitive, irreducible *qualitative* features of the physical universe. Such primitive qualities would exist over and above the elementary particles and fields that, according to contemporary physics, compose the furniture of the world and are exhaustively describable in quantitative terms.

But if we were thus to adorn the naked world of quantitative physics with color qualities, doing so would not seem to bring us any closer to answering the question of why the apple that you see looks red. The original quantitative scientific account tells us in a perfectly adequate way how the light with a particular spectral power distribution stimulates the receptors in the retina. We know that, with a properly functioning nervous system, this process is both necessary and sufficient for seeing object colors. Adding the supposed primitive qualitative redness of the apple to this account seems to play no additional role in helping us to understand how it is we come to see the apple as red.

Subjectivism

Confronted by these obstacles, we might consider abandoning our original supposition that the apparent redness of the apple is a property of the apple. Instead, we might think of the redness as our *subjective,* mental response to some feature of the physical apple. This suggestion goes back to the scientific revolution of the 17th century. It had been voiced some 2,000 years earlier by the Greek atomist Democritus, and repudiated by Aristotle, who preferred to locate colors outside of living creatures, in the interaction of light and matter. Medieval as well as ancient thinkers had followed Aristotle in this. For them, the qualities that objects are seen to have are by and large the qualities that they do have; these thinkers held that the business of our senses is to reveal Nature's finery rather than to clothe things with raiment of our own devising. Galileo introduced a new view of the Book of Nature: It is, he was fond of saying, written in the language of mathematics. The objects of mathematics are *quantitative:* number, structure, and motion. The powerful new mathematical physics seemed not to allow for the *qualitative* features of experience. Colors, sounds, and odors were thus swept from the physical world of matter in motion into the dustbin of the mind. There they joined thinking, purpose, and consciousness, which the new physics had likewise dispossessed from the world of matter. This sharp division of nature into a mental and a physical domain, famously advocated by René Descartes, served science well, but left obscure the nature of mind and its relationship to the physical world.

Although science and philosophy have become more sophisticated than they were in the 17th century, the basic picture of a divided nature remains with us. As some of its early critics remarked, this picture offends common sense. How could it be that the shapes that we see are inherent features of the physical world, whereas the colors that we see are constructions of the mind? Can we even imagine a shape with no color whatever? Furthermore, there are grave difficulties in understanding in what sense colors could be "in the mind." Colors are not, literally speaking, *in* minds as raisins are *in* cookies. We need more than a figure of speech. We need to know how color perception is to be understood if we are to think of colors as properties of mental items. Suppose that when you look at the apple that you take to be red, the direct object of your perception is an apple-shaped red color image. Your awareness of the redness of the apple—and, indeed of the apple itself—would be inferential awareness rather than the direct awareness that it seems to be. Worse still, because the physical apple is, by hypothesis, not colored at all, your supposed perception of its color is always illusory. But putting this aside, if mental objects are colored, we must be able to see them. And just how does that seeing take place? Is seeing a red image like seeing a red apple? If so, we will fall into an infinite regress because seeing an apple involves having a visual image of an apple, and if that is to be seen, there must be an image of that image, and so on. But if seeing a red image is not like seeing a red apple, then how do we see red images?

Epiphenomenalism, Eliminativism, and Mind and Body

If we suppose that there are neither red apples nor red mental images of apples, we must consider a more radical approach. Perhaps there are no color properties at all, just physical color stimuli on the one hand and mental color experiences (not red images, but experiences *as of* red) on the other. Taken together, these are sufficient to give us the appearance if not the reality of a world of colored objects. If we are willing to accept this position, we must still confront the most difficult problem of all: the relation of mental experiences to brain events. If experiences are identical to brain events, the brain as a physical entity must have qualitative as well as quantitative characteristics. And yet, qualitative characteristics have no more part to play in neuroscience than they do in physics. We might think of them as idle by-products, mere *epiphenomena,* of neural activity. If so, it is unclear how they could have come into being as a result of processes of natural selection because they can confer no advantage on the organism that possesses them. For this reason, we might be tempted to *eliminate* qualitative experiences altogether from our picture of the world, as previous generations of scientists eliminated other explanatorily useless items such as phlogiston and the electromagnetic ether. This would mean that we human beings are nothing more than biochemical computing machines. It would seem that to take this path is to turn our backs on our problem rather than solving it. We know what it would mean for a machine to act as if it perceives color, but it is hard to conceive how one could go about constructing a machine that actually experiences color (or feels pain, or falls in love). Would it make a difference if the machine were made of wetware rather than hardware? If so, why?

Our investigation of the place of color in the world has led us to confront the classic philosophical problem of the relationship between mind and body. One of the motivations of those who want to think of colors as being items in the physical world is to bypass, or even partially dissolve, this formidable issue. But even if this could be achieved, we would still have to deal with colored afterimages, color illusions, color experiences of migraine sufferers, and a whole host of chromatic phenomena that are not part of the world outside of our skins.

The problems we have touched upon are difficult and their solutions are evasive, but they cannot be brushed aside. The fact is that we are physical beings as well as beings that think, and feel, and see colors. There must be a unitary picture of ourselves and of our world that accommodates and explains this fact. It is up to us to construct that picture. Once we understand what it is to see the apple as red, we shall have gone a long way to understand what we are.

C. L. Hardin

See also Color Perception; Color Perception: Physiological; Computer Consciousness; Consciousness; Mary the Color Scientist; Mind and Body; Nonveridical Perception; Philosophical Approaches; Qualia

Further Readings

Byrne, A., & Hilbert, D. (1997). *Readings on color: Vol. 1. The philosophy of color.* Cambridge: MIT Press.

Hardin, C. L. (1993). *Color for philosophers: Unweaving the rainbow.* Indianapolis, IN: Hackett.

Matthen, M. (2005). *Seeing, doing, and knowing: A philosophical theory of sense perception.* Oxford, UK: Clarendon Press.

COLOR CONSTANCY

We would rather eat a banana that looks yellow than one that looks green because the banana's color appearance carries reliable information about its ripeness. In general, color appearance is a useful percept because it provides reliable information about object identity and state. When we search for our car in a large parking lot, we rely on color to pick out likely candidates; our driver's licenses list the color of our eyes and hair for identification; we detect that a friend is embarrassed by the blush of his or her face.

For color appearance to provide useful information about objects, it must correlate with properties intrinsic to objects and be stable against transient features of the environment in which the objects are viewed, such as the illumination. This stability, which is provided in good measure by our visual systems, is called *color constancy*. That we have generally good constancy is consistent with everyday experience. We are content to refer to objects as having a well-defined color, and only rarely (e.g., when looking for our car in a parking lot illuminated by sodium vapor lamps) do we observe large failures of constancy. This entry describes the problem of color constancy, empirical observations, and theories of constancy.

The Problem of Color Constancy

Vision obtains information about objects through the light reflected from them. If the reflected light were in one-to-one correspondence with physical object properties, then extracting stable object percepts would be straightforward. But the reflected light confounds properties of the illumination with those of the object. In the case of color, the relevant object property is its surface reflectance function $S(\lambda)$: the fraction of incident illuminant power that is returned to the eye at each wavelength λ. The relevant property of the illuminant is its spectral power distribution, $I(\lambda)$: the amount of power at each wavelength that arrives at the object. The spectrum reflected to the eye is thus $C(\lambda) = I(\lambda)\,S(\lambda)$. $C(\lambda)$ is called the color signal. Ambiguity arises because of the symmetric role played by $I(\lambda)$ and $S(\lambda)$ in the formation of the color signal. For example, a banana seen in skylight might reflect the same spectrum to the eye as does grass under direct sunlight because the effect of the illuminant change on the color signal can be counteracted by the change in surface reflectance. The perceptual challenge of color constancy is to make object color appearance stable against changes in $I(\lambda)$ while making it sensitive to changes in object reflectance $S(\lambda)$.

Empirical Observations

To what extent does the visual system actually stabilize object color in the face of illuminant changes? This has been studied with scaling and naming paradigms where observers describe the color appearance of objects seen under different illuminants, as well as with matching paradigms where observers adjust a test object seen under one illuminant to match the appearance of reference object seen under a different illuminant. A few generalizations may be drawn from the large empirical literature. First, color appearance does vary somewhat when the illuminant is changed: color constancy is not perfect. Second, the variation in object appearance is small compared with what would be predicted for a visual system with no constancy.

That constancy is generally good is often characterized by a constancy index, which takes on a value of 0 for a visual system with no constancy and 1 for a visual system with perfect constancy. For natural viewing conditions when only the illuminant is changed, experimentally measured constancy indices are often in the range 0.8 to 0.9, and sometimes higher.

Constancy is not always good, however. For example, constancy fails for simple scenes. Indeed, when a scene consists only of a single diffusely illuminated flat matte object, changes of illumination and reflectance are perfectly confounded, and changes of illumination lead to large failures of constancy. More generally, the degree to which the visual system exhibits constancy depends critically on what is varied in the scene. Under natural viewing conditions, constancy tends to be good if only the spectrum of the illuminant is varied. But if the surface reflectances of the other objects in the scene are covaried with the illuminant, constancy can be greatly reduced. For example, suppose the illuminant is shifted to have more short wavelength power and less long wavelength power. If at the same time, the surface reflectances of objects in the scene are shifted to compensate (i.e., to reflect less at short wavelengths and more at long wavelengths), color constancy is impaired. In laboratory studies of this manipulation, constancy indices drop into the range 0.2 to 0.4, a result that needs to be explained by any theory of constancy.

Theories

Theories of constancy should account for the general empirical observations described previously.

They should explain why constancy is often good, but also why it sometimes fails. Essentially all current theories share in common the general notion that visual processing of the color signal reflected from a single object is affected by the color signals reflected from the other objects in the scene. That is, our perception of object color is constructed by analyzing the reflected color signal relative to the rest of the retinal image.

Fundamentals of Color Vision

To understand theories of constancy, it is necessary to review a few fundamentals of human color vision. The color signal arriving at each retinal location is not represented completely. Rather, its spectrum is coded by the responses of three classes of light-sensitive photoreceptors. These are referred to as the L, M, and S cones, where the letters are mnemonics for long-wavelength-sensitive, middle-wavelength-sensitive, and short-wavelength-sensitive. Each cone class is characterized by a spectral sensitivity that relates the cone's response to the intensity of incident light at each wavelength. The three classes of cones differ in the region of the spectrum they are most sensitive to. Thus, the information about color available to the brain consists of the responses r_L, r_M, and r_S of the L, M, and S cones at each image location.

Contrast Coding

The simplest theories of constancy postulate that the initial representation of the image is processed separately for each cone class and that at each image location the cone responses are converted to a contrast representation. For the L-cones, contrast is based on the difference between the overall L-cone response r_L and the L-cone responses in its local neighborhood, and it expresses this difference relative to the magnitude of the neighboring responses. Let u_L represent the average of the L-cone responses in a spatial neighborhood of an L-cone whose response is r_L. Then the contrast is given by $c_L = (r_L – u_L)/u_L$. Parallel expressions apply for the M and S cones.

Experiments that assess the visual system's response to spots flashed against spatially uniform backgrounds support the idea that cone responses are converted to a contrast representation early in the visual system. These experiments include measurements of appearance, of visual discrimination thresholds, and direct measurements of electrical activity in retinal ganglion cells.

How does contrast coding help explain color constancy? Consider a change in the overall intensity of the illuminant spectrum. This will increase the cone responses equally to the light reflected from every location in the scene. Thus contrast, which depends on the ratio of cone responses across locations, is invariant to such illuminant changes.

When a change in the illuminant is not characterized as a simple scalar change in intensity, contrast is not guaranteed to be exactly invariant across the illuminant change. Nonetheless, analyses of the light reflected from natural surfaces indicate that contrast computed separately for each cone class (i.e., c_L, c_M, c_S) is quite stable across naturally occurring illuminant changes, even those that are not pure intensity changes. When the objects in the scene are held fixed and only the spectrum of the illuminant is varied, theories based on local contrast coding do a good job of accounting for experimental color constancy data.

Challenges for Contrast Coding Models

Although contrast coding can explain constancy across illuminant changes, it also predicts strong failures of constancy when the reflectances of some of the objects in the scene are changed. Suppose we fix one test object of interest, and then vary the objects around it at the same time as the illuminant is changed. Manipulations of this sort can be arranged so that the local means (u_L, u_M, u_S) near the test object remain fixed across the illuminant change. In turn, this means that the contrasts of the test object (c_L, c_M, c_S) vary with the illuminant, and contrast theories thus predict large failures of constancy under these conditions. Although experimental tests do reveal that color constancy is reduced when this type of manipulation is performed, the reduction is not as great as contrast theories predict. In addition, contrast is a quantity computed independently of the three-dimensional scene being viewed. Experimental manipulations that affect perceived depth relations without changing cone contrasts have been shown to affect how color is perceived. Contrast-based theories are incomplete.

Theoretical Approaches and Considerations

The current theoretical challenge is to develop models that can account for the full range of experimentally measured performance, when stimulus manipulations include changes in the illuminant, in the reflectances of other objects in the scene, and in geometric aspects of the scene.

One approach to model development is mechanistic. The idea underlying mechanistic models is that the cone responses corresponding to each object are subjected to a series of transformations as they propagate from the retina to the cortex, and that the end result of these transformations is a neural representation that is more directly correlated with object reflectance than are the cone responses. The contrast coding idea reviewed earlier is an example of a mechanistic approach, and contrast coding is generally taken as the first stage of modern mechanistic theories. Current research along mechanistic lines seeks to incorporate additional stages. For example, most models postulate that signals from the separate cone classes are recombined at second-stage opponent sites, which code differences between cone contrasts. Signals leaving the opponent sites are then further modified in an image-dependent fashion. For example, the overall magnitude of cone contrasts across the image are thought to control the gain applied to second-stage signals, and such processes have been shown to have the effect of further stabilizing color appearance across natural illuminant changes. To the extent that mechanistic models eventually account for color constancy across a wide range of experimental conditions, they have the attractive feature that they simultaneously provide an account of the chain of neural processing that underlies the color perception.

A second modeling approach is computational. Here the idea is to step back from the details of human vision and ask instead how to design an algorithm that, if it had access to the cone responses, could in principle achieve approximate color constancy. The algorithmic work lies within the domain of computer vision and has led to methods that explicitly estimate the physical scene illuminant $I(\lambda)$ and surface reflectances $S(\lambda)$ from the spatial array of cone responses. Although the resulting algorithms do not have any necessary connection to human vision, they may be elaborated into models that predict how object colors will be perceived. For example, one can link algorithm and human performance by predicting that two objects viewed in different scenes and under different illuminants, but whose surface reflectances are estimated by the algorithm to be the same, will also have the same color appearance. Modeling efforts of this sort have been successful at accounting for experimental measurements of object color appearance across a wide range of experimental conditions and particularly at accounting for both successes and failures of constancy in a unified fashion.

Computational modeling typically incorporates constraints derived from analysis of what surfaces and illuminants occur in natural scenes. In the Bayesian approach to developing computational methods, such constraints are expressed explicitly as probability distributions. This in turn allows a precise statistical formulation of how to optimally estimate surface and illuminant properties from the cone responses. In the context of developing models for human vision, the motivation for incorporating optimality principles into the development of algorithms is that evolution has driven the visual system toward information processing optimized for natural scenes, so that a computational method that is optimized with respect to the same set of scenes is likely to mimic human performance.

The mechanistic and computational approaches are complementary. Any computation that correctly predicts performance must have a mechanistic expression (because the human visual system implements it), and conversely, any mechanistic formulation that describes performance must express a computation that does an excellent job of achieving color constancy for natural viewing (because we are only rarely surprised to see a fixed object appear a different color.) The current challenge is to elaborate both of these approaches to describe the full range of performance, and to bring them together into a single model that connects mechanism (including contrast coding), performance, and a computational understanding.

David H. Brainard

See also Bayesian Approach; Color: Philosophical Issues; Color Naming; Color Perception; Color Perception: Physiological; Computational Approaches; Constancy; Lightness Constancy

Further Readings

Brainard, D. H. (2004). Color constancy. In L. Chalupa & J. Werner (Eds.), *The visual neurosciences* (Vol. 1, pp. 948–961). Cambridge: MIT Press.

Mausfeld, R., & Heyer, D. (Eds.). (2003). *Colour perception: Mind and the physical world.* Oxford, UK: Oxford University Press.

Shevell, S. K., & Kingdom, F. A. A. (2008). Color in complex scenes. *Annual Review of Psychology, 59,* 143–166.

Smithson, H. E. (2005). Sensory, computational, and cognitive components of human color constancy. *Philosophical Transactions of the Royal Society of London B, 360,* 1329–1346.

COLOR DEFICIENCY

Color vision is the capacity to distinguish objects based on the wavelengths of light they reflect or emit, and color vision is one of the most salient features used to describe objects in our environment. Consider a few things that exquisite color discrimination enables—being able to determine the ripeness of fruit, to rapidly detect your car in a parking lot, or to determine if your steak is cooked to your liking. Each individual's perception of color is a private sensation and, for those individuals who possess a "normal" ability to perceive color, this sense is often taken for granted. Conversely, individuals with color deficiency are often aware of their defective color discrimination throughout daily life. For example, people with color vision deficiencies can have difficulty telling whether a piece of meat is raw or well done, or might have difficulty seeing that they have been in the sun too long and are getting sunburned. They may also have a different sense of what looks good—be it their clothing or the food in front of them. This is because the color cues allowing individuals with normal color vision to distinguish between two objects may not be present for color-deficient individuals. For example, ketchup and chocolate syrup might look about the same color to them. This entry examines the range of color vision deficiencies present in humans, including their genetic origins, receptoral mechanisms, associated perceptual consequences, and incidence in the population.

Photoreceptor Mechanisms

The four types of photoreceptors in the human retina initiate the process of vision, and these types can be separated into two main classes (rods and cones). Rods are effective only at low light levels and can detect even a single photon of light, thus affording our scotopic vision (i.e., nighttime vision). Cones are less sensitive than rods are and therefore operate at higher light levels (serving photopic, or daytime, vision), and are responsible for our ability to perceive color. Thus, most of our visual experience is mediated by the cone photoreceptors (paradoxically, however, the cones constitute only about 5% of the total number of photoreceptors). There are three different classes of cone photoreceptor, each one with a different spectral sensitivity (i.e., plot of sensitivity of the photopigment vs. wavelength). Each of the four types of photoreceptor contains a single type of photopigment molecule, which gives the photoreceptor its characteristic spectral sensitivity.

As hypothesized 200 years ago by Thomas Young, we now know that trichromatic color vision is afforded by the presence of three spectrally distinct cone photopigments in the retina. Each cone type can be classified by the wavelength of light that the photopigment it contains is most sensitive to: short-, middle-, and long-wavelength sensitive (abbreviated S, M, and L). All humans with normal color vision have the same S-cone pigment, with peak absorption around 417 nanometers (nm). The M-cone pigment varies slightly between individuals, but peaks around 530 nm. Interestingly, there is widespread variation of the L-cone pigment among humans with normal color vision, but the peak of the L-cone pigment is approximately 559 nm. Rods are maximally sensitive to 500 nm light, but do not contribute significantly to our color vision.

The absorbance spectrum of a photopigment represents the probability of a photon of light being absorbed as a function of wavelength. In other words, each cone photoreceptor is simply a photon "counter"—an individual cone does not have access to information about the wavelength content of a given light stimulus. At the most basic level, the signal from a photoreceptor depends on the number of photons caught, not on their wavelength. For example, even though an L-cone is about two times more sensitive to 559-nm light than to 500-nm light, if the 500-nm light is more than twice as bright, it will elicit a higher quantal catch in the L-cone, and thus, a stronger neural response.

Consequently, the ability to discriminate colors arises from the neural channels that extract the *ratios* of the photon catches in the different spectral cone types. As such, possession of one cone type from each of the three spectrally distinct classes (S, M, L) provides the basis for normal color vision. However, if input from one or more of the cone classes is missing, the result is a compromised ability to make chromatic discriminations, that is, a color vision deficiency. Although these individuals are often called "color blind," in most cases, this is an inaccurate description of their color vision. With rare exception, these deficiencies do not involve a loss of sensitivity to all colors—as implied by the phrase *color blind;* rather, certain colors that can be easily discriminated by a "normal" observer look the same to the color-deficient observer.

Red-Green Color Vision Deficiencies

The most common form of inherited color vision deficiency is one that affects the red-green (L–M cone) system. Among individuals of Western European ancestry, about 7 to 10% of males have a red-green color vision defect. The incidence is lower among Africans as well as smaller isolated populations such as Fijian Islanders and the Inuit. The incidence in females is much lower (approximately 0.4%) because the defects are inherited as X-linked recessive traits, though approximately 15% of females are carriers of a red-green defect. The general genetic causes of red-green color vision deficiency involve a disruption of the L/M gene array on the X chromosome, which occurs either through a physical rearrangement in genes in which one type is missing or through the introduction of an inactivating mutation in one of the genes. In both cases, the result is an absence of function of either the L- or M-cones. This is because a cone cannot absorb light without a functional photopigment.

Within the red-green color vision defects, there are two broad subcategories—dichromacy and anomalous trichromacy (Figure 1). Red-green dichromats base their vision on an S pigment and a single L *or* M pigment. Individuals with an absence of L-cone function are said to have a *protan* defect, and individuals with an absence of M-cone function are said to have a *deutan* defect. Individuals with less severe forms of the red-green defects, known as anomalous trichromats, do not have the normal M or L photopigments; instead they base their vision on an S pigment and either two spectrally distinct M pigments (*protanomals*) or two spectrally distinct L pigments (*deuteranomals*). They are trichromats, in that they have three different pigments, but are different from normal (anomalous). Deuteranomalous defects are by far the most prevalent of any of the congenital color vision defects, affecting nearly 5% of men.

Perceptually, the absence of a cone type can have differing effects (see color insert, Figure 15). Individuals with a protan (L-cone) defect are less sensitive to light in the long-wavelength (red) portion of the spectrum. Therefore, the brightness of red, orange, and yellow are reduced compared with that of a normal observer. Furthermore, these individuals may have problems in distinguishing red from green, as well as difficulties differentiating a red hue from black. In contrast, individuals with deutan (M-cone) defects exhibit a reduction of sensitivity to colors in the green region of the spectrum. Deutans suffer similar hue discrimination problems as the protans, but without the abnormal dimming. These discrimination errors are exploited in the design of color vision tests; however, the real-world consequences of having certain red-green deficiencies are minimal. Some individuals go unaware of their defect until tested in a laboratory setting.

Tritan Color Vision Deficiencies

Tritan defects are an inherited autosomal dominant abnormality of S-cone function. Individuals with tritan defects base their vision on an L and an M pigment; they are dichromats. Unlike the red-green defects, there is no evidence for anomalous trichromats with tritan color vision defects. These defects occur in both males and females with equal frequency and are believed to affect approximately 1 in 10,000 people. Because the S-pigment gene is on chromosome 7 and humans are diploid (i.e., they have two copies of each autosome), each S-cone photoreceptor expresses the S-cone pigment genes from both copies of chromosome 7. Consequently, a defect in one S-pigment gene is sufficient to cause tritanopia,

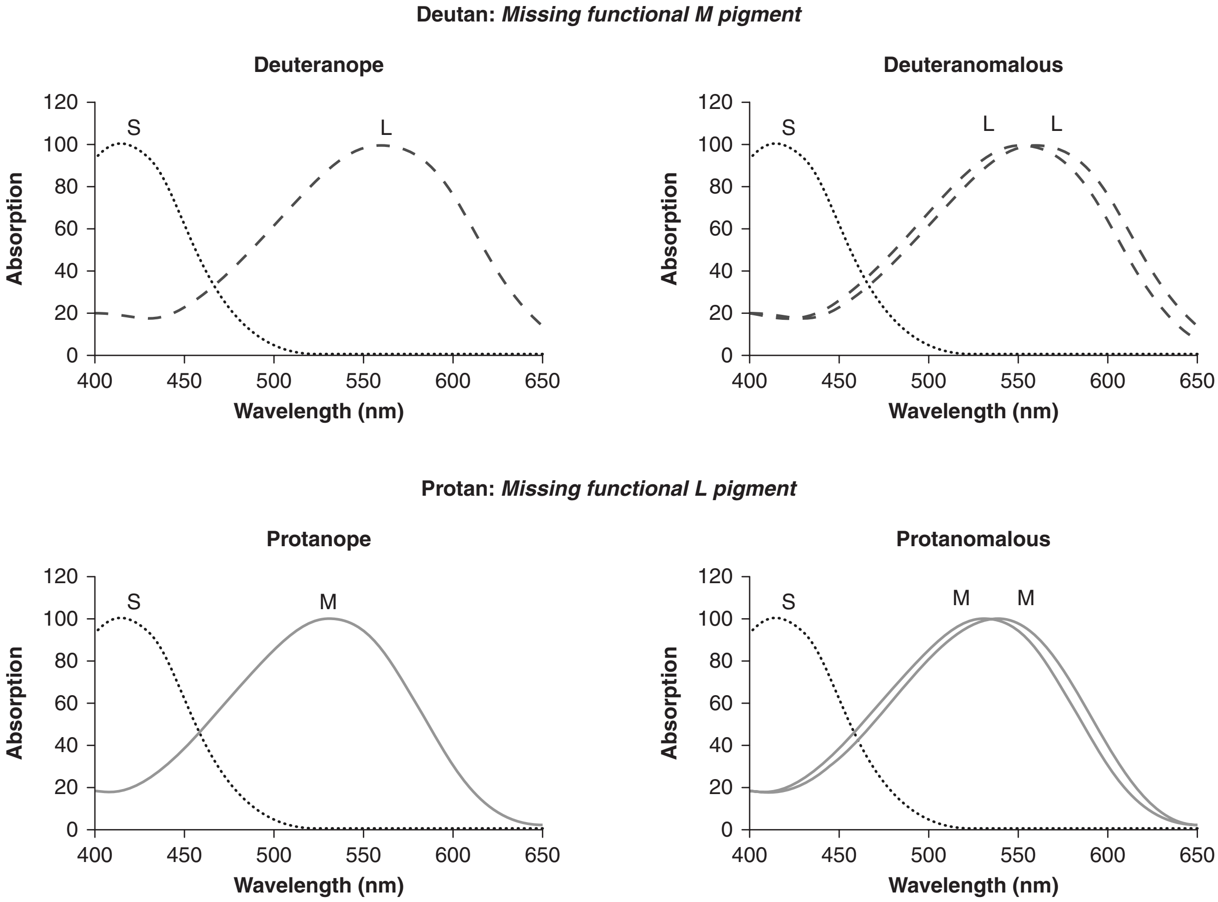


Figure 1 Photopigments Underlying Defective Red-Green Color Vision

Notes: Red-green color-deficient individuals are missing either all members of the M class or all members of the L class of pigment. Dichromats have only one pigment in the L/M region of the spectrum, whereas trichromats have two pigments in the L/M region of the spectrum. The degree of color vision deficiency in persons with anomalous trichromacy depends on the magnitude of the spectral difference between their pigment subtypes. Deuteranopes and deuteranomalous trichromats have no functional M pigment, though deuteranomals have two slightly different L pigments. Likewise, protanopes and protanomalous trichromats have no functional L pigment, though protanomals have two slightly different M pigments.

just as a mutation in one rhodopsin gene is sufficient to cause retinitis pigmentosa, a retinal degeneration that involves the degeneration of the rod photoreceptors. Furthermore, this disorder exhibits incomplete penetrance, meaning that individuals with the same underlying mutation can have different degrees of color-vision impairment, even between siblings.

Blue-Cone Monochromacy

Blue-cone monochromacy (BCM) is a condition where both L- and M-cone function is absent; therefore individuals will be able to function at normal light levels but will not be able to distinguish different hues. Any residual color vision in these individuals must be based solely on the S cones and rods. This means that under most conditions, they are monochromatic, though under mesopic (dim light; between photopic and scotopic levels) conditions, they may gain some additional discrimination capacity. This disorder affects approximately 1 in 100,000 individuals.

As with the red-green defects, there are two main genetic causes of BCM, sometimes referred to as one-step or two-step mutations, though both lead to the complete absence of functional L and M photopigment. In one-step mutations, there is an inability to express the L and M pigment genes—as such, no L or M photopigment is made. In the two-step mutations, the L or M pigment genes can be expressed; however, the pigment that is made is nonfunctional because of some other mutation. In both the one-step and two-step mechanisms, individuals have poor visual acuity, myopia, nystagmus (involuntary eye oscillations), and minimally detectable electroretinogram (ERG) responses. Because of the X-linked nature of the condition, female carriers are spared from a full manifestation of the associated defects, but they can show abnormal ERG amplitudes.

Achromatopsia

Complete achromatopsia (i.e., rod monochromacy) is a congenital vision disorder in which all cone function is absent or severely diminished. This disorder affects as many as 1 in 30,000 people and is thought to result from mutations in one of two components of the cone phototransduction cascade (transducin or the cyclic-nucleotide gated channel). Rod monochromacy is typically characterized by a complete lack of color discrimination, photophobia (light sensitivity), reduced acuity, visual nystagmus, and a nondetectable cone ERG. Because these individuals lack all cone function, perception is dominated by the rod system and vision at normal light levels can be difficult.

Acquired Defects

As described earlier, most color deficiencies are inherited and result from disruptions in the genes encoding the cone photopigments. However, color vision deficiencies may be acquired through exposure to toxins (such as heavy metals), cortical defects (occurring from a stroke or shaken baby syndrome), or other disease (diabetes, optic neuropathy, glaucoma). As a generalization, these defects are predominantly tritan, have an equal incidence in males and females, and are typically accompanied by a reduction in visual acuity. Acquired defects can be categorized by the perceptual color vision deficit or by the mechanism. Perceptual categorization can be broken into three basic categories: a nonspecific loss of color vision, a loss pertaining to the red-green axis, or a loss corresponding to the blue-yellow axis. There are also three mechanistic categories: absorption, alteration, and reduction. Absorption defects affect the absorption of a particular hue. For example, the normal aging process affects the lens of the eye, causing a selective absorption defect in discrimination in the blue-yellow axis. An alteration defect is one that shifts the normal color match, and normally reflects a destruction of the macular cones. A reduction defect occurs typically as a result of diseases of the optic nerve and causes perceptive reduction of saturation within the red-green axis as well as a milder blue-yellow loss.

Testing and Assessment

When screening for color vision deficiency and attempting to make diagnoses, no one test should stand in isolation. Over the last few centuries, several different methods for diagnosis and testing of color vision deficiencies have been developed, and using a battery of tests will help the examiner to assign a more accurate diagnosis. Furthermore, the use of a standard illuminant ensures reliability and repeatability of the test(s).

The anomaloscope, discovered by Lord Rayleigh in 1881, is a color-matching test in which subjects view a small bipartite field and compare a mixture of red (670 nm) and green (546 nm) wavelength light on one half to a spectral yellow (589 nm) on the other half. The *Rayleigh match* takes advantage of the fact that normal trichromats behave dichromatically for lights with wavelengths longer than 550 nm (because the S cones are not sensitive at these wavelengths). Thus, by restricting the lights employed to the longer wavelengths, a normal trichromat can achieve a perfect color match to the monochromatic light by mixing two "primaries." A subject's *match range* can be any range of numbers between 0 and 73, with most subjects matching two to four of the numbers in the range of 38 to 42. Individuals do vary in the relative amounts of the primaries needed to achieve the match, altering their match range slightly (this variability has been linked to individual variation in the L- and M-cone pigments). However, individuals with missing or altered L or M pigment will match a much larger and distinctly displaced range of mixtures of the two primaries because of their relative insensitivity to one of the wavelengths of light. Therefore, the extent of the matching range is directly representative of the severity of the defect. For example, deuteranomals need less red than normal in the red-green mixture to match the spectral yellow (matching a wide range of numbers in the 0–40 range), whereas protanomals need more red in the match (matching a wide range of numbers typically in the 40–73 range). A subject matching the entire 0–73 range is diagnosed as dichromatic. This test is considered by many to be the gold standard in diagnosing type and severity of red-green defects.

Because the Rayleigh match uses wavelengths to which the S cones are relatively insensitive, it only can be used to detect red-green defects (defects associated with disrupted L- or M-cone function). To evaluate S cone function, which provides the basis for color vision along the blue-yellow axis, one can use a color matching test called the *Moreland match*. The test field consists of a cyan (480 nm) and yellow (580 nm) light, and the matching primaries are blue (440 nm) and green (488 nm). This color match is often employed in the evaluation of acquired color vision because they more often affect the blue-yellow system.

In the interest of speed and convenience, printed pseudoisochromatic plate tests are a widely used screening tool, of which there are several varieties. Mostly these tests specialize in detecting red-green deficiencies; however, a few include plates to detect tritan defects. These tests operate on the pseudoisochromatic principle by using a field of dots that vary in luminance, and embedded within the dots is a number or other recognizable shape differing in chromaticity from the background. Four different types of plates can be included. Vanishing plates contain a figure that is correctly recognized by a color normal observer, but is not recognized by a color defective observer because the figure is confused with the background. The diagnostic plates are more sophisticated in that, based on the type of error, they can differentiate between a protan and deutan observer. One of the more advanced plate types is the transformation plate in which two figures are embedded: one figure recognized by a normal observer, and the other recognized by a color defective observer. Finally, in the hidden plate, the figure is seen by the color defective observer rather than by the normal observer. Current versions of these tests include Ishihara, AOHRR plates, and the Neitz Test of Color Vision.

Another type of test is the arrangement test in which subjects order materials in a specific hue or saturation order. The stimulus is prearranged in a random order and subjects are instructed to arrange the stimuli according to color. Correct arrangement of the various color samples requires normal color vision, as well as good discriminative ability. Some versions are so difficult that even individuals with trichromatic color vision can make significant errors. These tests vary in the number of colored samples used in the test; current versions of these tests include Farnsworth D15, Lanthony's desaturated D15, and the Farnsworth-Munsell 100 hue test.

An emerging tool for assessment of color vision is molecular genetic analysis. From a small blood sample, it is possible to isolate a person's DNA, which can be used to analyze the genes encoding the S, M, and L photopigments. The presence of mutations or deletions in one or more of these genes provides predictive information relevant to the subjects' color vision capacity. This means that in nearly every individual with a color vision deficiency, it is possible to determine a precise genetic cause for the deficiency.

Diane M. Tait and Joseph Carroll

See also Animal Color Vision; Color: Genetics of; Color Perception: Physiological

Further Readings

Birch, J. (1993). *Diagnosis of defective colour vision.* New York: Oxford University Press.

Gegenfurtner, K. R., & Sharpe, L. T. (1999). *Color vision: From genes to perception.* New York: Cambridge University Press.

Mollon, J. D., Pokorny, J., & Knoblauch, K. (2003). *Normal and defective colour vision.* Oxford, UK: Oxford University Press.

Rodieck, R. W. (1998). *The first steps in seeing.* Sunderland, MA: Sinauer Associates.

COLOR MIXING

Differently colored lights can be mixed together optically to produce lights of another color. Which lights and which mixtures perceptually match each other can be predicted from a few relatively simple rules. The simplicity of these rules, discussed in this entry, derives from the properties of the visual photoreceptors in our eyes that convert arriving photons into neural signals, rather than from physics.

Overview

The visible spectrum is that part of the electromagnetic spectrum that we can see. It covers wavelengths from about 390 to 730 nanometers (nm). When viewed alone, the appearance of *monochromatic* lights made up of single wavelengths varies across the spectrum from violet at short wavelengths through blue, blue-green or cyan, green, yellow-green, yellow, orange to red at long wavelengths. Sodium street lighting, which appears yellow, is a commonly encountered example of lights that are nearly monochromatic. Most lights that we encounter in the natural environment are *broadband* in the sense that they consist of lights of many different wavelengths covering broad regions of the spectrum.

Mixtures of monochromatic lights or mixtures of broadband lights can be perceived as identical even when the components in the mixtures are physically different. For example, a mixture of red and green monochromatic lights can appear identical to a yellow monochromatic light, and a mixture of blue and green monochromatic lights can appear identical to a cyan monochromatic light (see color insert, Figure 16, left panel). The relationship between the physical characteristics of lights or mixtures of lights and whether or not they appear to match can be investigated in simple color matching experiments.

Properties of Color Mixing

In a typical color matching experiment, an observer looks at a circle, half of which is illuminated by a "test" light of variable wavelength (λ) and the other half by a mixture of three "primary" lights usually chosen to look red (R), green (G), and blue or violet (B). For each test light, the observer adjusts the intensities of the three primary lights, so that the test field is perfectly matched by the mixture of primary lights. With the proviso that sometimes one of the primaries must be added to the test light to complete the match and that the primaries must be independent (in the sense that none of the primaries can be matched by a mixture of the other two), only three primaries are required to match any test light. The upper panel of Figure 1 shows the mean red, green, and blue color matching functions (CMFs) for primary lights of 645, 526, and 444 nm (denoted by the vertical dashed lines) as a function of test wavelength, λ. Each function defines the amount of that primary required to match a monochromatic light of λ nm. That any light can be matched by a mixture of just three primary lights reflects a fundamental property of normal human vision: it is *trichromatic*. The operation of color television, computer monitors, and projection systems, which produce colors by mixing together red-, green- and blue-appearing lights in different proportions (as you can see for yourself by looking closely at the screens), relies on our color vision being trichromatic.

Another important property of color mixing is that it behaves linearly as reflected in a set of descriptive laws known as Grassmann's laws. Color matches are symmetric (if light A matches light B, then B matches A), transitive (if A matches B, and B matches C, then A matches C), obey the proportionality rule (if A matches B, then n × A matches n × B), and are additive (if A matches B and C matches D, then the combination of A and C matches the combination of B and D).

The Biological Basis of Color Mixing

The origins of trichromacy and the linearity of color mixing are mainly biological. They occur because there are only three classes of cone photoreceptor in the human eye, each of which responds *univariantly* to light. Once a photon is absorbed, its effect is the same whatever its wavelength, as a result of which photoreceptors are, in effect, sophisticated photon counters.

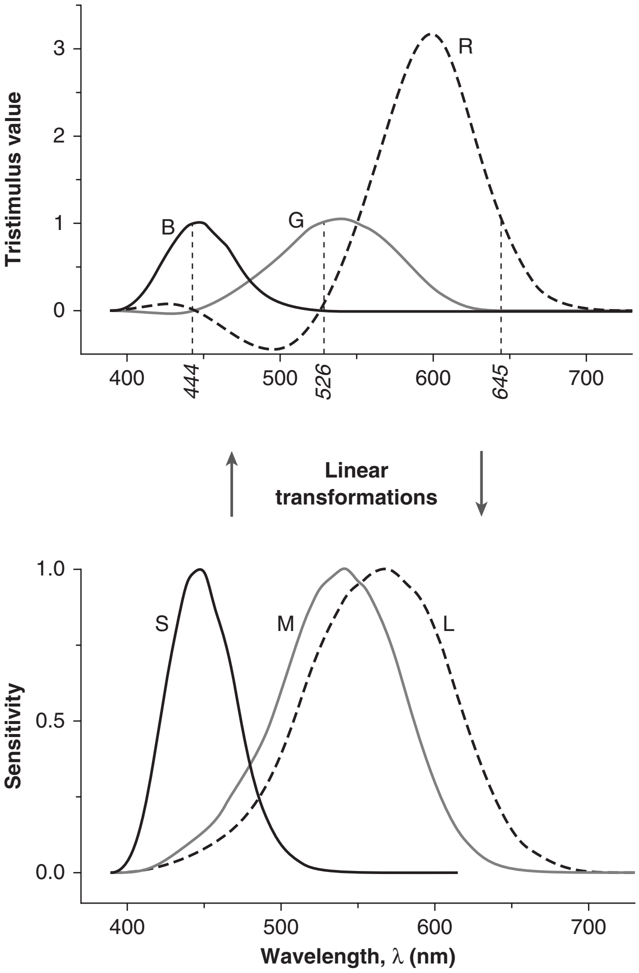


Figure 1 Color Matching Functions and Cone Spectral Sensitivities

Notes: The red (R), green (G), and blue (B) CMFs (upper panel) can be linearly transformed to the long-wavelength (L), middle-wavelength (M) and short-wavelength (S) cone spectral sensitivities and vice versa. The CMFs for primaries of 444, 526, and 645 nm are from Stiles and Burch, and the cone sensitivities are their linear transformation proposed by Stockman and Sharpe. By convention, CMFs are defined for test lights of equal energy across the spectrum.

Photoreceptors, however, are more likely to absorb some wavelengths than others (see Figure 1, bottom panel). A change in photoreceptor output could result from a change in wavelength, from a change in light intensity, or from both, so individual photoreceptors are color blind. Our ability to perceive color depends on the comparison between the outputs of the three cone photoreceptor types.

When observers match the test and mixture fields in a color matching experiment, the two fields are matched for each of the three cone types. In other words, the two lights that perceptually match generate the same pattern of cone outputs. Matches depend on the three cone spectral sensitivities. As illustrated in Figure 1, the red (R), green (G), and blue (B) CMFs (upper panel; shown as dashed, gray, and black lines here) are a linear transformation of the spectral sensitivities of the long- (L), middle- (M), and short- (S) wavelength-sensitive cone spectral sensitivities (lower panel).

Color Mixing and Color Appearance

Cone spectral sensitivities and CMFs tell us which pairs of lights match, but they do not tell us what the color of the matched pair looks like to the observer; that is, they do not tell us about color appearance. Color appearance depends also on the context within which the lights are presented. For example, lights that are identical can look different when surrounded by areas of different color. The effect on color appearance of changing the color of a surrounding area from blue to yellow is shown in the color insert, Figure 17. The two squares are printed with the same ink, but the square with the yellow surround looks much bluer than the square with the blue surround does. Color matches reflect trichromatic univariance at the cone photoreceptors, whereas color appearance depends on previous light exposure and, in addition, on the complex activity of post-receptoral mechanisms in the retina and brain that act on the outputs of the cone photoreceptors.

Additive and Subtractive Color Mixing

What this entry has discussed so far is the *additive* mixing of visually superimposed colored lights. Additive color mixing produced by mixing lights is illustrated in the left panel of Figure 16 in the color insert. In contrast, the mixing of paints, dyes, and pigments depends on the lights reflected by pigments (i.e., on the wavelengths that are not absorbed or “subtracted” by the pigment) rather than on the addition of projected lights. Thus, a yellow pigment appears yellow because it absorbs short-wavelengths and reflects middle- and long-wavelengths, and a cyan pigment appears cyan because it absorbs long-wavelengths and reflects short- and middle-wavelengths. When predicting the effects of mixing pigments, the lights that are subtracted by the mixture pigments must be considered. For example, a mixture of yellow and cyan pigments will appear green, because the two pigments absorb, respectively, short- and long-wavelengths, leaving only green-appearing, middle-wavelength light to be reflected. Subtractive color mixing is illustrated in the right panel of Figure 16 of the color insert.

Andrew Stockman

See also Color Deficiency; Color Perception; Color Perception: Physiological

Further Readings

Berns, R. S. (2000). *Billmeyer and Saltzman's principles of color technology* (3rd ed.). New York: Wiley.

Brainard, D. H. (1995). Colorimetry. In M. Bass (Ed.), *OSA handbook of optics* (Vol. 1, pp. 26.21–26.54). New York: McGraw-Hill.

Krantz, D. H. (1975). Color measurement and color theory: I. Representation theorem for Grassmann structures. *Journal of Mathematical Psychology, 12,* 283–303.

Sharpe, L. T., Stockman, A., Jägle, H., & Nathans, J. (1999). Opsin genes, cone photopigments, color vision and colorblindness. In K. Gegenfurtner & L. T. Sharpe (Eds.), *Color vision: From genes to perception* (pp. 3–51). Cambridge, UK: Cambridge University Press.

Stiles, W. S., & Burch, J. M. (1959). NPL colour-matching investigation: Final report (1958). *Optica Acta, 6,* 1–26.

Stockman, A. (2003). Colorimetry. In T. G. Brown, K. Creath, H. Kogelnik, M. A. Kriss, J. Schmit, & M. J. Weber (Eds.), *The optics encyclopedia: Basic foundations and practical applications* (Vol. 1, pp. 207–226). Berlin: Wiley-VCH.

Stockman, A., & Sharpe, L. T. (1999). Cone spectral sensitivities and color matching. In K. Gegenfurtner & L. T. Sharpe (Eds.), *Color vision: From genes to perception* (pp. 53–87). Cambridge, UK: Cambridge University Press.

Stockman, A., & Sharpe, L. T. (2000). Spectral sensitivities of the middle- and long-wavelength sensitive cones derived from measurements in observers of known genotype. *Vision Research, 40,* 1711–1737.

Wyszecki, G., & Stiles, W. S. (1982). *Color science: Concepts and methods, quantitative data and formulae* (2nd ed.). New York: Wiley.

COLOR NAMING

Color naming in the broad sense is giving a color name such as "red" to an object of a particular color. This task is not as trivial as it sounds. We commonly use only a few color names but can distinguish millions of different colors. The colors of an object such as a banana may differ to some extent in all perceptual color dimensions of hue, saturation, and brightness: Parts of the banana may be darker and less saturated because of shading, parts may be more yellow-greenish or more yellow-orangish. Yet we perceive all these different shades of colors as belonging to the same category "yellow" and speak about the "yellow" banana. Color naming thus involves a partitioning of the continuous color space into distinct color categories, to which we assign a word. In the narrower sense, color naming refers to a method used in psychological experiments where observers have to name the color of the stimuli presented. This entry describes the universality of color names, psychophysical experiments, the Stroop effect, and neural correlates.

Universality of Color Names

Color names in different languages are remarkably consistent. It has been proposed that 11 basic color terms are strikingly similar across different languages and cultures. The English words for these basic color terms are *black, white, red, yellow, green, blue, gray, orange, brown, pink,* and *purple.* These color names may be grouped and ordered to approximately reflect their appearance during child development and cultural development as follows: (black, white) → (red) → (yellow, green) → (blue) → (gray, orange, brown, pink, and purple). An analysis of data from 110 languages of nonindustrialized societies gathered in the World Color Survey confirmed the idea of universal color names. Although there are some small variations of the focal colors (those colors that provide the best example of a basic color term) across languages, the focal color fell on average close to the colors found for English speakers. Despite these similarities, different languages may differ in the number of basic color terms. For example, some languages such as Russian, Polish, and Italian have a basic color term to distinguish between "light blue" and "dark blue"; some languages merge green and blue into a single "grue" category, and members of the Dani of New Guinea have only two basic color terms to distinguish between light-warm and dark-cool colors. Interestingly, when asked to remember colors, the Dani were better across the same color boundaries (e.g., blue-green) than English observers were, even though their language does not have any corresponding color terms. Another New Guinea language, Berinmo, deviates from the universal pattern: Berinmo has a "nol"/"wor" category that goes right across the green category of English speakers. Overall, the issue of the universality of color names and categories is still controversial.

Psychophysical Experiments

Color naming used as a method in psychophysical experiments gives a direct report of the color sensation evoked by the stimulus. Color naming seems to be an inherently coarse method because there are only a few color names but millions of distinguishable colors. One way to overcome this limitation is to use a technique called *hue scaling,* where observers rate the amount of red, green, blue, and yellow in the stimulus. The shapes of the hue scaling curves have been found to be highly similar across observers, and only the amplitude differed. Further, observers almost never rated a color as being both red and green, or both yellow and blue, in agreement with the notion that red-green and blue-yellow are opponent colors. Another way to overcome the limitations of color naming is to let observers name a large number of color samples and to infer the partitioning of color space based on the naming results of all color samples arranged on the color plane according to the chromaticities of the color samples. This method has been used to investigate color constancy, showing that observers can achieve almost perfect color constancy (the ability to see objects in the same color, independent of the illumination).

The Stroop Effect

The naming of a color interferes with the reading of a word (the so-called Stroop effect): For example, people take longer to name the blue ink color of the printed word *RED* than if the blue ink is used to print the word *BLUE.* In general, the naming of the ink color of a printed word is harder if the word is a color name that is different from the ink color. The Stroop effect occurs even though the observers were instructed to ignore the letters and to pay attention only to the ink. The Stroop effect shows that color naming is tightly coupled to reading, another learned linguistic skill. The Stroop effect occurs also when observers have to name the color of objects with a characteristic object color shown in a "wrong" color. In the reverse Stroop task, observers are instructed to ignore the ink color in which a color word is printed and to respond to the meaning of the color word. Again, reaction times are generally faster with congruent combinations than with incongruent combinations.

Neural Correlates

The neural correlates of the color categories have not yet been found and are a topic of intense research. The present results support the view that in early visual areas V1 to V4 basic color categories do not have a special status. A potential candidate area for the representation of color categories is an area in the inferior temporal (IT) cortex.

Thorsten Hansen

See also Color Perception; Cultural Effects on Visual Perception

Further Readings

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution.* Berkeley: University of California Press.

Hansen, T., & Gegenfurtner, K. R. (2006). Color scaling of discs and natural objects at different luminance levels. *Visual Neuroscience, 23*(3–4), 603–610.

Lindsey, D. T., & Brown, A. M. (2006). Universality of color names. *Proceedings of the National Academy of Sciences USA, 103*(44), 16608–16613.

Naor-Raz, G., Tarr, M. J., & Kersten, D. (2003). Is color an intrinsic property of object representation? *Perception, 32*(6), 667–680.

Regier, T., Kay, P., & Cook, S. R. (2007). Focal colors are universal after all. *Proceedings of the National Academy of Sciences USA, 102*(23), 8386–8391.

Smithson, H. E., Khan, S. S., Sharpe, L. T., & Stockman, A. (2006). Transitions between color categories mapped with a reverse Stroop task. *Visual Neuroscience, 23*(3–4), 453–460.

COLOR PERCEPTION

Color vision is defined as the ability to discriminate differences in the wavelengths of light. Yet, like all vision, the *perception* of color is ultimately not about seeing wavelength or light but about seeing the objects that light illuminates. Color provides powerful cues to object recognition. For example, it can help reveal what sort of surface we are viewing (e.g., leaf or bark) and the particular state that the material is in (e.g., dry or lush foliage, ripe or unripe fruit, or a pallid or glowing complexion). To make these judgments, the color percept must be tied to properties of the object (how it reflects light) rather than simply to the light reaching the eye (which can change as the illumination changes). The strategies and extent to which the visual system achieves this perceptual goal is the subject of color constancy.

Color can tell us what we are looking at and where it is by helping to segment different parts of the scene based on their colors (see color insert, Figure 18). Black and white photography shows that color is not necessary for seeing the layout of a scene, and many studies have found that color cues on their own do not support the perception of fine spatial details nor as strong a sense of motion as luminance cues. Indeed, the chromatic equivalent of a black and white image—where only differences in color but not in luminance are shown—is often unrecognizable. Yet color is nevertheless one of the most salient cues for segregating objects. A ripe fruit pops out from green leaves, and the leaves in turn appear grouped together to form trees, because differences in color set objects apart whereas similarities in color bind them. Thus, color vision plays an important role in spatial vision. Chromatic cues may be especially important for seeing shapes in the presence of shadows. Cast shadows produce much larger changes in brightness than in color, and thus the ability to link together parts of the scene that have a common color allows the underlying objects to be delineated more clearly. Individuals with color deficiencies often find spatial tasks, where items must be located or grouped by their color, most difficult. This entry discusses color appearance and contextual influences on color appearance

Color Appearance

The seemingly limitless variety of colors we can experience (estimates suggest that humans can discriminate among millions of different shades) rests on detecting the responses from only three types of cone photoreceptors with different sensitivities to wavelength. This trichromatic representation fundamentally determines what information the brain can extract about a light's spectrum, and it is probably not coincidental that the perceptual representation of color is itself three-dimensional. The principal psychological attributes of a color are its hue (e.g., whether it is red or blue), saturation (the degree to which it appears mixed with white, as in pink versus red), and brightness (the perceived intensity). Perceptual color spaces attempt to capture the relationships between different colors—typically in terms of these attributes—and these arrangements are thought to reveal how color percepts are built by the visual system.

In modern color spaces, hues are arranged around a circle with white at the center, an idea that may have first appeared in 1704 with the publication of Sir Isaac Newton's *Opticks*. In this color circle, different hues thus correspond to different directions from white, while saturation is captured as the distance from the achromatic center (see color insert, Figure 19). If these circular planes are stacked according to their brightness from dark to light, then the space becomes a three-dimensional color solid such as a cone or sphere. The actual structure of the solid can be complicated because of limits imposed by the realizable gamut of colors and because the perceptual distances between colors can vary in complex ways. A further complication is that the different attributes of color do not vary independently. For example, the hue of many wavelengths changes when they are desaturated by adding white light (a phenomenon known as the Abney effect), or when their intensity is varied (the Bezold-Brucke effect). Specifically, the hues of wavelengths tend to become more blue or yellow as intensity increases.

Among all the hues around the circle, some appear special or more fundamental. Experiments with color appearance suggest that red, green, blue, and yellow seem psychologically pure, in that they cannot be reduced to more basic sensations. In contrast, all other hues appear to be mixtures of two of these primaries. For example, orange appears to contain both red and yellow, and purple is a mixture of red and blue. A striking feature of color experience is that some mixtures are never perceived together. We cannot (under simple viewing conditions) perceive a color that is both red and green, or both blue and yellow, in the same way that we cannot see a light that is both bright and dark. Such observations led the 19th-century physiologist Ewald Hering to propose the opponent process theory of color vision. This held that color is represented by three mechanisms that signal red versus green, blue versus yellow, or bright versus dark. Because red and green are represented by the same mechanism responding in opposite ways (e.g., exciting to red and inhibiting to green), the red and green primaries are mutually exclusive sensations, but can co-occur with the sensations carried by the mechanism opposing blue and yellow.

By this model, the sensation of a pure red (one not tinged by blue or yellow) occurs at the unique point where the blue-yellow response is zero. In turn, a pure gray is perceived when both the red-green and blue-yellow responses are zero. The stimuli corresponding to the unique hues of red, green, blue, and yellow have been studied extensively because they help define the responses of the opponent processes to different wavelengths (their spectral sensitivities). Three of the unique hues can be seen at wavelengths within the rainbow spectrum of visible light. Unique blue is roughly at 470 nanometers (nm), green at 510 nm, and yellow at 580. (However, as noted later, the actual locations vary widely among individuals.) Unique red is instead known as an extraspectral light. The longest wavelengths we can see appear too orange (red tinged with yellow), and thus, a pure red requires mixing these wavelengths with a bluish, short wavelength to cancel out the yellowness. More generally, the yellowness at any wavelength can be cancelled by adding blue light, whereas redness can be cancelled by adding green light. This hue cancellation technique was used by the psychologists Leo Hurvich and Dorothea Jameson beginning in the 1950s to map out the sensitivities of the red-green and blue-yellow responses to different wavelengths, work that prompted a modern revival of Hering's opponent process theory.

A different approach to studying how the spectrum of hues is parsed was developed by the anthropologists Brent Berlin and Paul Kay in the 1960s. They asked observers to label the colors falling on the surface of a perceptual color space (the Munsell color solid) and to pick out the chip that provided the best example of a color label. Berlin and Kay's work was designed to ask how color naming varies across speakers of different languages. This method defines a set of primaries corresponding to basic color terms—terms that are used reliably by most speakers of a language and that refer to colors independent of objects (e.g., unlike salmon or lilac). The English language has 11 basic terms, including the Hering landmarks (red, green, blue, yellow, black, white, and gray) and additionally brown, orange, purple, and pink. This suggests a more complex representation of color than would be conveyed by independent responses along the three opponent dimensions. For example, pink and red may represent separate focal points in color appearance even though they are both shades of red. Moreover, the color categories of yellow and brown occur only for light or dark colors respectively, and thus require an interaction between the Hering's red-green, blue-yellow, and light-dark dimensions.

In the late 1950s, researchers including Russell De Valois and his colleagues began recording the color responses of single cells in the visual pathway of primates. Many cells in the eye's retina and in the lateral geniculate nucleus to which retinal cells project gave responses of opposite polarity to different parts of the spectrum—exciting to some wavelengths while inhibiting to others. This provided a physiological basis for color opponency. These post-receptoral cells could build opponency simply by receiving inputs of opposite sign from the different classes of cones, similar to the way that spatial opponency was known to be built by receiving signals of opposite sign from the center and surround of the cell's receptive field. This helped reinforce opponent color theory, which remains a cornerstone of modern color science.

The actual links between color appearance and color coding have remained elusive. Specifically, it remains unclear how color sensations are represented by the responses of visual neurons. Among the two principal types of cells responding to color in the geniculate, one responds best to color changes that vary between red and blue-green, and the second signals color changes that perceptually vary from purple to yellow-green. Thus, the perceptually pure axes of red-green and blue-yellow cannot be tied to a single cell type at this level of the visual system. Moreover, the basis for unique hues becomes even less clear when signals from the geniculate are passed on to the visual cortex, where cells are found tuned to many different color directions. Whether there is a subset of cells or a transformation at further stages in the visual pathway that culminate in neurons with the responses predicted by perceptual opponency remains an area of active interest and debate.

Because of such discrepancies, many color scientists have also begun to look for the unique hues in another place—outside the observer. Some colors may look more special because they correspond to special properties of the environment rather than to special states in the neural response. For example, the phases of natural daylight vary along a blue-yellow axis, and thus unique blue and yellow could reflect learning about a prominent physical characteristic of the world. Similarly, red (and pink) might be special because they are the ecologically important colors of blood and skin tones and are colors that signal ripening fruit among the foliage. The idea that color appearance is determined by the color environment has been bolstered by the findings that the wavelengths corresponding to the unique hues cannot be predicted from the visual sensitivity of observers. For example, humans vary dramatically in the relative number of different cone types, yet this variation is not correlated with observers' choices for unique hues. Large individual differences in sensitivity also occur because of differences in the amount of inert pigments in the eye screening the receptors. The lens contains a yellow pigment that steadily increases with age, blocking progressively more of the short wavelength light reaching the retina. Yet, the unique hues and the white point remain stable across age, suggesting that color appearance is continuously recalibrated to compensate for sensitivity changes in the observer.

Finally, all accounts of color appearance—physiological, environmental, and cultural—are plagued by the problem that individual differences in the unique hues and in color naming are enormous, even among individuals living in a common environment and speaking a common language. This raises the possibility that the perceptual representation of color is not strongly determined by any factor. Why people describe colors the way that they do thus remains an enigma.

Contextual Influences on Color Appearance

The dimensions of hue, saturation, and brightness are adequate to describe color appearance under simple conditions, when observers view a single uniform spot in an otherwise dark room. Percepts in this viewing mode are known as unrelated colors and appear to describe attributes of light. Yet, when stimuli are instead judged in the presence of other colors and in spatially varying contexts—as is typically the case in natural viewing—then the experience of color becomes much richer and more complex. These related colors more closely reflect how we see colors in the real world and point to the importance of perceiving color as a property of objects and surfaces.

A clear example of the perceptual changes that can emerge from this change in viewing context is the distinction between brightness and lightness. As noted, brightness refers to the perceived intensity of light coming from a spot, whereas lightness corresponds to the perceived reflectance of the spot—what percentage of the incident illumination is returned from the spot. The paper on this page has a high reflectance and appears white, whereas the print absorbs most of the light and appears black. Note these are perceptions of the surfaces themselves and not simply of the amount of light at a point on the page. Whether the page is viewed in a dim room or bright sunlight—an intensity change that can be many orders of magnitude—the print and paper still appear black and white, and the reason the print looks dark is that it reflects less of the incident light than the paper does. If a spot from the print could be viewed through a tiny aperture that left the rest of the scene in darkness, then the spot would suddenly appear bright, a phenomenon known as the Gelb effect. The change occurs because the print is now the only point with light. Any light on a completely dark background is an increment and will be perceived as light. Dark colors can only be perceived when there are decrements (i.e., when the background is brighter). As a result, the color brown can never be seen as an unrelated color, and only arises when there are other elements in the scene that are brighter.

How do surrounding colors affect color appearance? One process is simultaneous contrast or induction. When a spot is placed on a uniform surround, the perceived lightness depends roughly on the ratio of the center and surround intensities rather than the absolute intensities (see color insert, Figure 20a). Thus, the same gray appears darker on a bright surround but brighter on a dark surround. Perceived hue is similarly biased, so that a gray becomes greenish or reddish when placed on a red or green background. Context can also bias color appearance in the opposite way, by reducing the perceived color differences between surfaces in the scene. These effects are known as assimilation, and typically occur in finely varying patterns where the colors begin to blend. Finally, color appearance can be affected by the average color in the background and by its variance. For example, a color will appear less saturated when the gamut of surrounding colors varies widely, even though the average color in the surround remains neutral.

For simple patterns, these contrast effects depend largely on the local differences between the center and surround, and are an example of a general principle that the visual system tends to emphasize relative differences in stimuli. However, the contextual effects may also reflect much more sophisticated inferences about the stimuli. A number of recent (as well as some classic) illusions have shown that subtle changes in the spatial arrangement of patterns can lead to dramatic changes in the perceived lightness and color of the elements. These changes tend to be those that alter the perception of transparency or shading, suggesting that the perception of lightness and hue includes computations that factor in the three-dimensional structure and illumination in the scene. One striking illustration of these effects is neon color spreading, where the perceived color of line elements tends to diffuse into the surround areas to give the impression that the lines are being seen through a colored filter.

Context can also affect color over time rather than over space. Like all sensory systems, the visual system is highly adaptable and adjusts its responses whenever the stimulus changes. These adaptation effects are well known in color vision. To a person who has just viewed a red square, a gray square appears greenish. This occurs because the visual system becomes less sensitive to the adapting color, thus biasing the response to other stimuli away from red and toward its opposite, green. A second effect is that the red itself becomes less saturated and thus more neutral. With careful fixation, a blurry blob of color will completely fade to gray (see color insert, Figure 20b). This adaptation re-centers color space so that the white point corresponds to the average color we are exposed to, discounting the changes in color appearance that would otherwise result from a change in the average color of the lighting. Many of the contextual influences in color appearance can be seen as processes that are designed to compensate for variations in the environment or in the observer so that the perceptions of color remain faithful to the properties of surfaces.

As these examples illustrate, in natural viewing situations, the perceived color of a patch cannot easily be predicted simply from the light spectrum because it also depends on the surrounding colors and context. Yet even these contextual influences fail to capture the full richness of color perception. Real surfaces have many perceptual qualities and can, for example, appear rough or smooth, or lustrous or dull. These attributes cannot be reproduced by uniform spots in surrounds because they depend on the spatial and spectral variations within rather than between surfaces. The surface statistics that the visual system uses to judge the actual material properties of surfaces is currently an active area of research.

Michael Webster

See also Color Constancy; Color Naming; Color Perception: Physiological; Context Effects in Perception; Contrast Perception; Cultural Effects on Visual Perception; Individual Differences in Perception; Lightness Constancy

Further Readings

Abramov, I., & Gordon, J. (1994). Color appearance: On seeing red—or yellow, or green, or blue. *Annual Review of Psychology, 45,* 451–485.

Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., 339–351). Cambridge: MIT Press.

Fairchild, M. D. (2005). *Color appearance models* (2nd ed.). Chichester, UK: Wiley-IS&T.

Gegenfurtner, K. R., & Sharpe, L. T. (Eds.). (1999). *Color vision: From genes to perception.* Cambridge, UK: Cambridge University Press.

Kaiser, P., & Boynton, R. M. B. (1996). *Human color vision.* Washington, DC: Optical Society of America.

Kuehni, R. G., & Schwarz, A. (2008). *Color ordered.* Oxford, UK: Oxford University Press.

Mausfeld, R., & Heyer, D. (Eds.). (2004). *Colour perception: Mind and the physical world.* Oxford, UK: Oxford University Press.

Monnier, P., & Shevell, S. K. (2003). Large shifts in color appearance observed with patterned backgrounds. *Nature Neuroscience, 6,* 801–802.

Shevell, S. K. (Ed.). (2003). *The science of color.* Oxford, UK: Elsevier.

Shevell, S. K., & Kingdom, F. A. A. (2008). Color in complex scenes. *Annual Review of Psychology, 59,* 143–166.

Webster, M. A., & Kay, P. (2007). Individual and population differences in focal colors. In R. E. MacLaury, G. V. Paramei, & D. Dedrick (Eds.), *Anthropology of color* (pp. 29–53). Amsterdam: John Benjamins.

Werner, J. S. (1998). Aging through the eyes of Monet. In W. G. R. Backhaus, R. Kliegl, & J. S. Werner (Eds.), *Color vision: Perspectives from multiple disciplines* (pp. 3–41). Berlin: De Gruyter.

Werner, J. S., Pinna, B., & Spillmann, L. (2007). The brain and the world of illusory color. *Scientific American, 296,* 90–95.

COLOR PERCEPTION: PHYSIOLOGICAL

When we view the world around us, we attribute color to the individual objects we look at. But, as was discovered by Sir Isaac Newton, the light rays being reflected from these objects into our eyes are not colored. How we see color depends on the visual processing going on in the eyes and the cortex. Color vision starts with the absorption of light in three different types of light sensitive receptors in the eye, which convert electromagnetic energy into electrical signals, which in turn are transformed into action potentials by a complicated network of cells in the retina. The information is sent to the visual cortex via three independent channels with different chromatic preferences. In the cortex, information from these channels is mixed to enable perception of a large variety of different hues. Furthermore, recent evidence suggests that color analysis and coding cannot be separated from the analysis and coding of other visual attributes such as form and motion. Although some brain areas are more sensitive to color than others are, color vision emerges through the combined activity of neurons in many different areas. This entry describes the physiological mechanisms of color perception.

Cone Receptors

Human color vision is characterized by three different classes of cone photoreceptors, denoted long- (L), middle- (M), and short-wavelength-sensitive- (S) cones. Light is absorbed in the cones and converted by a complex photochemical reaction into an electrical signal, whose magnitude is determined by the number of photons absorbed by each cone. Figure 1 shows the most recent and precise estimates of the cone absorption spectra. Thus, although the light stimulus is determined by its intensity at an infinite number of wavelengths between approximately 400 and 700 nm, the output of the cones can be characterized by only three numbers—the absorption for each cone class. This is the principle of *trichromacy*. The three classes of cone therefore drastically reduce the dimensionality of color vision. Several characteristics about the human cone absorption spectra are worth noting. Figure 1 shows that all types of cone are sensitive to a wide range of wavelengths. L- and M-cones are sensitive nearly over the whole visible spectrum. The L-cones have their peak sensitivity at a wavelength that would appear yellowish and the S cones have their peak sensitivity at a wavelength that would appear violet under neutral viewing conditions. Therefore, the common labeling of the cones as red, green, and blue is misleading because these labels do not correspond to the wavelength to which the cones are maximally responsive. Thus, there is no simple relationship between the excitation of a single class of photoreceptors and the color we perceive.

Instead, photons of different wavelengths have different likelihoods of being absorbed by the three cone classes. Once absorbed, the only remaining information is the photon count in each cone, not the wavelength of the absorbed photons, a principle termed *univariance*. An increase in the photon count of a photoreceptor can thus result from an increase in light intensity, a change to a more favorable wavelength for that cone class, or both. Therefore, to compute the color of an object unambiguously, the magnitudes of the output signals of the three cone classes have to be compared. This starts at the

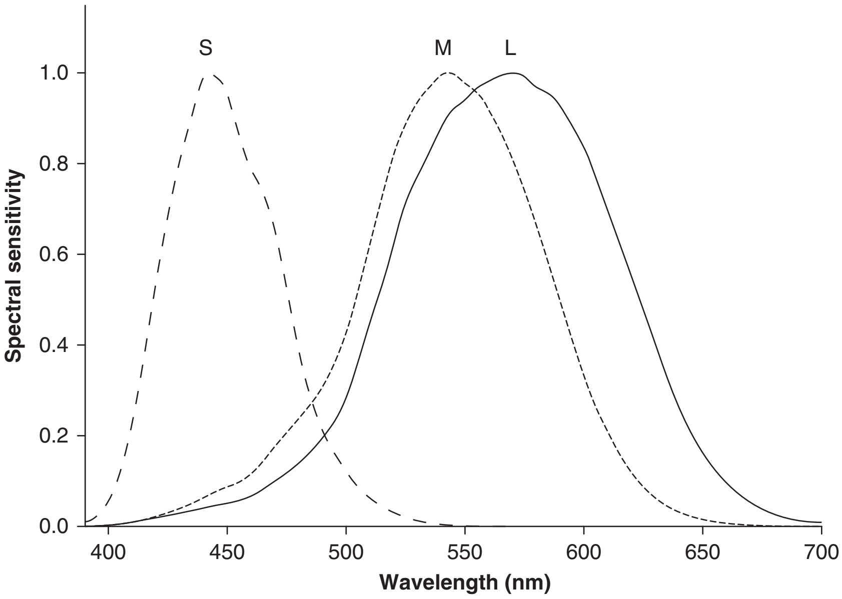


Figure 1 Cone Absorption Spectra

Notes: The relative sensitivities of the short- (S), middle- (M), and long- (L) wavelength sensitive cones as a function of wavelength. Each curve was normalized to its maximum.

next stage of processing, which is performed by bipolar cells and ganglion cells in the retina.

Retinal Ganglion Cells and Lateral Geniculate Nucleus

Three channels of information from the retina to cortex are distinct from each other in their chromatic properties, as well as in their anatomical substrate. These three channels do not simply correspond to the signals conveyed by the three different classes of photoreceptors. Instead, the signals from the cones are repackaged into three new types of signals. It is assumed that the major reason for this transformation is the high degree of correlation in the signals from the L- and the M-cones, which results from the high degree of overlap in their spectral absorption functions.

One of these channels is called the luminance channel because it simply computes the intensity of stimulation by taking a weighted sum of the outputs of the three different cone types. This channel does not convey information about the color of things. The other two channels are called cone-opponent because they take differences of the output signals of the three different cone types. Ganglion cells in all three channels have circularly symmetric receptive fields with excitatory and inhibitory subregions.

There is one type of ganglion cells in the retina, called M-cells as in magnocellular, with relatively large dendrites. The M-cells do respond to luminance only, similar to the way a black-and-white printer would represent an image by taking a weighted mixture of signals from L- and M-cones. Interestingly, the weights that have been found physiologically by recording from individual M-cells are identical to the weights determined for photometric luminance in flicker-photometric experiments on human observers. About half of M-cells have a region at the center of their receptive field where they receive excitatory input from a group of L- and M-cones, whereas the surrounding parts of the receptive field receives inhibitory inputs. These are called ON-center cells. The other half has the opposite receptive field arrangement and therefore responds best to dark spots on a white background. These cells are called OFF-center cells.

More interesting from the viewpoint of color vision are the other two types of cells. Ganglion cells that contribute to the P-channel are called parvocellular and have smaller dendritic fields. In and near the fovea, the centers of their receptive fields are formed by single L- or M-cones, making these ganglion cells suitable for high acuity vision. The receptive field surround is formed mostly by the other cone class, that is, if there is an L-cone in the center, the surround would be dominated by M-cones. As with the M-cells, center and surround have opposing signs. For the P-cells, this has two consequences. First, if the stimulus is a white spot on a dark background, ON-center P-cells would respond quite well to it. Second, if the input is a large long-wavelength uniform field of light, an excitatory center with an L-cone would be maximally stimulated, and the inhibitory surround with M-cones would be maximally suppressed, leading to a maximal activation of the cell. Therefore, P-cells respond best to structured luminance stimuli and to uniform colored stimuli. They are said to fulfill a "double-duty" because they potentially underlie both color vision and high acuity luminance vision. Note, however, that P-cells do not respond to structured color stimuli, such as a small spot of red surrounded by a green of the same luminance. Actually, P-cells will not respond at all to such a stimulus because the excitation at the center is matched in that case by the inhibition in the surround.

Finally, it was recently discovered that there is another color opponent channel to the cortex, the so-called koniocellular (K) pathway that carries signals from S-cones to a large degree. These ganglion cells take the difference of the signals from S-cones and a summed signal of the L- and M-cones, and these cells are usually not spatially opponent. Therefore, these cells respond to large blueish or yellowish fields of light, but not to luminance patterns.

Thus, three channels of information from the retina to cortex are distinct from each other both in their chromatic properties and in their anatomical substrates. The differences in the chromatic properties of P- and K-cells is shown in Figure 2(a), which shows a distinct clustering of the chromatic preferences of lateral geniculate nucleus (LGN) cells along the L–M, and the S–(L+M) axes of color space. When these cone-opponent mechanisms were first discovered in the monkey retina, it was assumed that these would be the neural substrate of

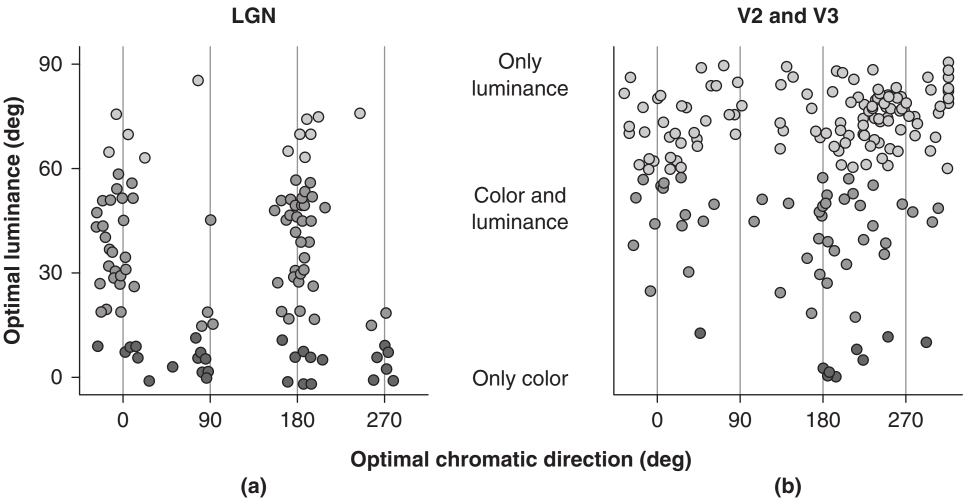


Figure 2 Color and Luminance Preferences of Neurons in the LGN and in the Cortex (V2 and V3)

Notes: The *x*-axis denotes the optimal chromatic direction; the *y*-axis denotes the preferred luminance as elevation above the isoluminant plane. Neurons in the LGN cluster around the cardinal directions, neurons in the cortex do not. Most neurons in the LGN and the cortex respond to both luminance and color.

the color-opponent mechanisms that Ewald Hering had proposed at the beginning of the 20th century. Hering had defined his "unique hues" purely phenomenally by observing that a color perceived as red never contained green as well, even though it could be a blueish red (purple) or a yellowish red (orange), and vice versa. Analogously, blue would never appear yellow. These opponent colors strongly suggested a link with the cone-opponent retinal mechanisms. However, many years of research showed that these two concepts are clearly distinct. For example, the S–(L+M) retinal cone-opponent mechanisms respond best to colors that appear purple and greenish-yellow, but not so much to colors that we would typically call good examples of blue and yellow, such as in the Swedish flag.

Hering's unique hues are formed at the level of cortical color processing, presumably at the stage where we group different colors into categories. For example, when we view a rainbow, we do not perceive a gradual transition from one wavelength to the next, but we see a typical stripe pattern. The physiological basis for this categorization process is still unclear. Several high-level visual cortical areas of the monkey brain have been suggested as the neural basis, but at this time, it is not even clear that monkeys have the same color categories and unique hues as humans.

Processing of Color in the Early Visual Cortex

Initial studies of the functional architecture of the primate visual cortex found only a small percentage of about 10% of all neurons devoted to the analysis of color signals. In these studies, the quest was for neurons that would only respond to stimuli defined by color, and give no response to luminance. According to a view promoting a functional segregation, it was thought that these neurons would analyze the color of objects, and other specialized neurons would deal with form, motion, and depth. In contrast to these results, later neuroimaging studies observed a strong and vigorous response to color that exceeded the response to luminance stimuli. The solution to this seeming contradiction was found when the color properties of single neurons in the primary visual cortex were investigated in more detail. A large proportion of neurons responds well both to color and to luminance. This is in line with anatomical findings that the three streams that arrive in the visual cortex from the retina do not remain segregated but become mixed to a large degree. Along these lines, many recent studies failed to find a strong segregation of functional properties in the primary and secondary visual cortex. Rather, it seems that the tuning of each cell for different visual stimulus attributes is more or less statistically independent. For example, knowing that a cell is selective for color does not allow a prediction about the orientation selectivity of that cell.

Although it was initially thought that the chromatic properties of V1 cortical cells are quite similar to those of the retinal cells, more details have been discovered in recent years. In the retina, responses of a neuron to all colors can be predicted just by the knowledge of the cells' preferred color because these neurons sum their cone inputs in a linear manner, and the preferred colors fall into two clusters of L–M and S–(L+M). Although most cells in V1 still follow the linear model, the color preferences are more widely distributed, as shown in Figure 2(b). These cells still show quite a broad tuning to color, responding to many different hues. In V2, a subpopulation of neurons was found that has more restrictive responses limited to a rather narrow range of hues, mimicking the behavior of higher order color mechanisms observed psychophysically. Similar effects have been reported in V4, and there seems to be a progression from V1 via V2 to V4 and the inferotemporal (IT) cortex in the specificity of the chromatic response.

Another important characteristic of color cells that emerges in V1 is double opponency. As mentioned earlier, the single opponent cells in the retina and LGN cannot signal spatial and chromatic properties at the same time. They respond to color without considering the spatial context. Double opponent cells are both spatially and chromatically opponent, and are therefore able to signal chromatic boundaries—for example, a red fruit next to a green leaf. These cells effectively compute color contrast—the difference between two colors, so their response would be independent of the average color in the scene, which is mostly determined by the illumination color. Therefore, this would be an ideal mechanism to achieve color constancy, which is the ability of the primate visual system to assign a stable color to an object, independently of how it is illuminated. Only recently have these double-opponent cells been found in the primate's visual cortex. They are members of the class of cells that respond equally well to luminance and color.

Evidence for Specialized Cortical Regions

Early evidence for the existence of a unique color center in the visual cortex came from single-unit recordings of macaque monkey area V4. It was argued that area V4 is uniquely specialized for color because it contains an unusually large proportion of color selective cells, some of which exhibit the property of color constancy. Initial estimates of the proportion of color selective cells in V4 were as high as 100%. Subsequent studies found significantly lower estimates, some as low as 20%. Although a consensual estimate is still lacking, it is likely that V4 contains approximately the same proportion of color selective cells as do V1, V2, or V3. Moreover, the color selectivity of V4 cells appears to be similar to that of neurons in earlier areas. Most V4 neurons are not more narrowly tuned in their wavelength selectivity than retinal ganglion or LGN cells. However, a subpopulation of V4 neurons exists that is more narrowly tuned. In that respect, V4 is not different from V1 or V2.

The major argument against V4 as the color center of the monkey brain comes from lesion studies. Lesions of extrastriate area V4 do lead to mild deficits in color vision and to a variety of other visual deficits. Monkeys with V4 lesions are severely impaired in shape discrimination, object recognition, texture discrimination, and in their ability to focus attention. Lesions to the next processing stage, the inferotemporal cortex, seem to mimic the human condition of cerebral achromatopsia, that is, a specific loss of color perception. However, the color deficit depends on the removal of the entire IT cortex, which has a variety of other dramatic effects on higher-level vision. In the IT cortex, the proportion of color selective cells is believed to be high as well—somewhere between 50% and 70%. The distribution of preferred colors in the IT cortex was reported to be rather uniform, with a subpopulation of cells being selective to a narrow range of colors. These results suggest that the IT cortex could be involved in the elaboration of perceptual color categories.

In humans, it has long been reported that bilateral cortical damage to the lingual and fusiform gyri on the ventromedial side of the occipitotemporal lobe could lead to severe color vision deficiencies. In analogy to earlier single-unit studies in monkey V4, this area has been called human V4 (hV4). Brain imaging studies show this part of the visual cortex to contain a representation of the contralateral hemifield and to be highly active during a variety of color vision tasks.

Although all these results imply that some region of the visual cortex, whether it is V4 or IT, has a high sensitivity to color relative to luminance, no single experiment has shown evidence for a region that does respond to color *only*. Similarly, the visual world of patients with cerebral achromatopsia is not the world of black-and-white movies. The patients typically show a variety of severe object- and pattern-recognition disorders. Some researchers think it more likely that color perception is the result of the simultaneous activity of neurons belonging to several cortical areas. That different aspects of color perception can be differentially affected by cortical lesions supports this view. Indeed, several patients have been reported with lesions outside the primary visual cortex; they had impaired color constancy but no deficits in color discrimination. Similar results had been reported for monkeys with V4 lesions. In addition, color-memory and color-constancy can be differentially affected in humans suffering from cortical lesions. Thus, as is the case for most other visual attributes, our experience of color probably depends on the activity of neurons belonging to a number of different cortical areas. This notion is further supported by the observation that most visual cortical areas are activated by exposure to chromatic stimuli, and in area hV4, the relative response to color might just be slightly higher than in other areas of the brain.

Karl Gegenfurtner

See also Color Constancy; Color Deficiency; Color Perception; Neural Representation/Coding; Retinal Anatomy; Visual Processing: Retinal

Further Readings

Bouvier, S. E., & Engel, S. A. (2006). Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cerebral Cortex, 16*(2), 183–191.

Callaway, E. M. (2005). Structure and function of parallel pathways in the primate early visual system. *Journal of Physiology, 566*(Pt 1), 13–19.

Dacey, D. M., & Packer, O. S. (2003). Colour coding in the primate retina: Diverse cell types and cone-specific circuitry. *Current Opinion in Neurobiology, 13*(4), 421–427.

Gegenfurtner, K. R. (2003). Cortical mechanisms of colour vision. *Nature Reviews. Neuroscience, 4*(7), 563–572.

Gegenfurtner, K. R., & Kiper, D. C. (2003). Color vision. *Annual Review of Neuroscience, 26,* 181–206.

Sincich, L. C., & Horton, J. C. (2005). The circuitry of V1 and V2: Integration of color, form, and motion. *Annual Review of Neuroscience, 28,* 303–326.

Solomon, S. G., & Lennie, P. (2007). The machinery of colour vision. *Nature Reviews. Neuroscience, 8*(4), 276–286.

Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron, 56*(2), 366–383.

COMMON CHEMICAL SENSE (CHEMESTHESIS)

In 1912, G. H. Parker described the sensory system responsible for detecting chemical irritants as "the common chemical sense." Although Parker noted that the common chemical sense was mediated by irritant-detecting free nerve endings resembling pain receptors, he concluded that it was a separate sense with its own set of receptors that produced distinct sensation qualities different from taste and smell. Free nerve endings are axons that do not terminate in specialized structures and are devoid of myelin (fatty insulation around axons). Today, we know that the nerves that respond to chemical irritants do not constitute a separate, independent sense, as Parker hypothesized. Rather, they are part of the general somatic sensory system: a subset of pain and temperature sensitive fibers that can be found throughout the skin and mucosal membranes of the nose, mouth, respiratory tract, eye, and anal and genital orifices. Recently, the term *chemesthesis* has been used to describe the sensations elicited by the chemical stimulation of free nerve endings. Chemesthesis is thus analogous to somesthesis, which describes sensations produced by mechanical stimulation of the general somatic sensory system. We stimulate our chemesthetic sense every time we eat chili peppers, sniff ammonia, or cut onions. These same nerve endings respond to other environmental irritants that may affect our health and determine the food we eat and the personal products we buy. This entry describes trigeminal chemoreception, stimulation of nerve endings, solitary chemoreceptor cells, and reversal of aversion.

Trigeminal Chemoreception

In humans, chemesthesis is best represented by free nerve endings in the trigeminal (Vth cranial) nerve innervating the mouth, nose, and eyes. However, free nerve endings originating from other cranial nerves, such as the glossopharyngeal (IX) and vagus (X) in the oral cavity and throat, as well as spinal nerves innervating the skin, also respond to chemical stimuli. A great deal of the information on chemical irritation has been obtained from cells in the dorsal root ganglia of the spinal cord that give rise to the spinal nerves. Nevertheless, much of the research on chemesthesis as it relates to environmental irritants and especially taste and smell involves trigeminal chemoreception. One main reason is that trigeminal free nerve endings are accessible to chemicals, lying in areas that are directly exposed to the environment.

The trigeminal nerve is a mixed sensory nerve containing axons of different sizes. The cell bodies of these axons are found in the trigeminal ganglia situated near the base of the brainstem. Trigeminal nerve fibers have the ability to respond to touch, cold, heat, and painful stimuli. The fibers are of different diameters and conduct action potentials at different speeds. The largest fibers respond to touch and conduct action potentials rapidly. The smallest fibers (unmyelinated and small myelinated) respond to painful stimuli and conduct action potentials slowly. Fibers that respond to chemical irritants are polymodal—they respond to a variety of noxious, irritating stimuli (mechanical, thermal, and chemical)—and are especially sensitive to capsaicin, the active, "hot" ingredient in chili peppers.

Stimulation of Nerve Endings

A variety of sensations are elicited by stimulation of trigeminal chemesthetic nerve fibers, including those described as pungent, tingling, stinging, burning, cooling, warming, painful, and irritating. In addition to providing a sensory experience, chemical stimulation of trigeminal chemoreceptors activates protective reflexes, such as increasing secretions (e.g., saliva, tears, and nasal mucus), decreasing breathing, initiating sweating, and decreasing the size of the nasal passages. Anyone who has bitten into a chili pepper or taken a whiff of ammonia has experienced some of these reactions. These reflexes function to counteract the effect of the irritant by diluting it or preventing more of it from entering the body.

Numerous chemical irritants stimulate trigeminal free nerve endings. Some chemesthetic stimuli may affect human health and welfare by eliciting unpleasant sensations in the eyes, nose and throat and potentially harmful physiologic reactions (including respiratory and cardiovascular changes, and allergic reactions). Many indoor air contaminants fall into this category and may contribute to building-related illness or sick building syndrome. These compounds are called volatile organic compounds (VOCs). Some VOCs include nicotine and benzene (tobacco smoke), formaldehyde (pressed wood products), limonene (aroma and cleaning products), and toluene and xylenes (craft and building supplies).

Some of the best-studied irritants that elicit responses from chemesthetic trigeminal nerves are the pungent compounds found in a wide variety of plants. Pungent vegetables and spices contain irritating compounds such as allyl isothiocyanate (horseradish and mustard), eugenol (cloves), gingerol and zingerone (ginger), piperine (black pepper), menthol (mint), cinnamaldehyde (cinnamon), propanethial-S-oxide (onions), and capsaicin (chili pepper).

How these chemicals activate trigeminal nerve endings is actively being explored in many laboratories. Although the mechanism of stimulation is known for only a few compounds, researchers believe irritants interact with receptor proteins found on the chemesthetic free nerve endings. Receptor proteins are like locks that are opened by appropriate chemical stimuli, the keys. Once the locks are opened, ions (charged atoms) move into the nerve endings and initiate signals (action potentials) that are carried to the brain to be deciphered. Some of the best studied of these receptor proteins are activated by nicotine and acidic stimuli (low pH). The carbon dioxide in carbonated beverages is irritating, in part, because it undergoes a reaction that results in a low pH.

In the mid 1990s, a new class of receptor proteins that is activated by irritants was discovered. These proteins belong to the transient receptor potential (TRP) channel family. The first of these proteins was reported by David Julius and his colleagues at the University of California, San Francisco. This polymodal receptor protein, now called TRPV1, is activated by stimuli such as heat, low pH, piperine, eugenol, and capsaicin. Other TRP channels have now been discovered that are triggered by compounds such as allyl isothionate, cinnamaldehyde (TRPA1), and menthol (TRPM8).

Many chemesthetic stimuli activate olfactory and taste receptor cells at low concentrations and are non-irritating. At higher concentrations, they elicit responses from chemesthetic trigeminal nerve fibers in the mouth, nose, and eyes and are described as irritating. For example, although benzaldehyde (which smells like almonds) and NaCl (table salt) are pleasant at low concentrations, both can become irritating at high concentrations. Researchers can determine at what concentration odorous compounds stimulate nasal trigeminal nerves by measuring the lateralization threshold—the concentration of the compound that can be detected in one side of the nose. For example, if a low concentration of an odorant such as benzaldehyde is delivered to one nostril and clean air to the other nostril, the nostril that received the odorant cannot be determined. However, once the concentration of the odorant becomes high enough to stimulate trigeminal receptors, the correct nostril can be resolved. That is, any concentration of the irritant that can be localized to one of the nostrils must be stimulating trigeminal nerves. Knowing the lateralization threshold has allowed researchers to make psychophysical measurements of chemesthesis.

Stimulation of nasal and oral trigeminal chemoreceptors may affect the perception of odorants and tastants. For example, chemical trigeminal stimulation could limit the access of odorant molecules to the olfactory organ reflexively by modifying breathing, nasal secretions, or the size of the nasal passages. Activation of nasal trigeminal nerve fibers by irritants may also directly inhibit olfactory receptor neurons by axon reflex. Cell bodies in the trigeminal ganglia send their axons to the nasal cavity where they may make several branches. One branch may go to the respiratory epithelium, but another may go to the olfactory epithelium housing the olfactory receptor neurons. Once an action potential is produced in the respiratory epithelium branch of the nerve, it travels to the brain and to the olfactory epithelium branch. Here, substances may be released that decrease the activity of olfactory receptor neurons. Similarly, evidence indicates that stimulation of oral trigeminal nerves may indirectly or directly decrease responses to taste stimuli. The purpose of this arrangement may be to keep high concentrations of chemical irritants from overwhelming the more sensitive olfactory and gustatory systems.

Solitary Chemoreceptor Cells

As described earlier, chemesthesis traditionally has been thought to involve only trigeminal free nerve endings in the nasal and oral cavities and the eyes. Recently, however, cells have been discovered that may change that idea. These cells, called solitary chemoreceptor cells (SCCs) are found scattered throughout the nasal cavity and the respiratory and digestive tracts. Many of these cells have the characteristics of individual taste cells that detect bitter compounds. The SCCs in the nasal cavity are contacted by trigeminal nerve fibers. Bitter substances activate nasal trigeminal nerve fibers, presumably through the SCCs. Other compounds may also stimulate trigeminal nerve fibers through the SCCs.

Reversal of Aversion

An interesting aspect of chemesthesis is that many substances that are initially (and probably innately) aversive eventually become preferred. Examples include the consumption of such substances as tobacco, chili pepper, mustard, curry, horseradish, ginger, and vinegar. Infants, children, and naive adults typically reject these stimuli on first exposure. However, most adults reverse their natural rejections and acquire strong positive preferences for at least one innately unpalatable substance. Indeed, whole cuisines, such as Mexican and Korean, are based on some of these chemesthetic stimuli. Why the reversal of the initial aversion occurs is not well understood.

Wayne L. Silver

See also Animal Chemical Sensitivity; Cutaneous Perception; Olfactory Localization; Olfactory Receptors and Transduction; Pain: Physiological Mechanisms; Psychophysical Approach; Taste Receptors and Transduction; Temperature Perception

Further Readings

Bautista, D. M., Jordt, S. E., Nikai, T., Tsuruda, P. R., Read, A. J., Poblete, J., et al. (2006). TRPA1 mediates the inflammatory actions of environmental irritants and proalgesic agents. *Cell, 124,* 1269–1282.

Bryant, B., & Silver, W. L. (2000). Chemesthesis: The common chemical sense. In T. E. Finger, W. L. Silver, & D. Restrepo. (Eds.), *Neurobiology of taste and smell* (2nd ed., pp. 73–100). New York: Wiley-Liss.

Cain, W. S., Lee, N. S., Wise, P. M., Schmidt, R., Ahn, B. H., Cometto-Muñiz, J. E., & Abraham, M. H. (2006). Chemesthesis from volatile organic compounds: Psychophysical and neural responses. *Physiology and Behavior, 88,* 317–324.

Finger, T. E., Bottger, B., Hansen, A., Anderson, K. T., Alimohammadi, H., & Silver, W. L. (2003). Solitary chemoreceptor cells in the nasal cavity serve as sentinels of respiration. *Proceedings of the National Academy of Sciences USA, 100,* 8981–8986.

Julius, D., & Basbaum, A. I. (2001). Molecular mechanisms of nociception. *Nature, 413,* 203–210.

COMPUTATIONAL APPROACHES

Perception is the analysis of sensory input in the context of our prior perceptual experience of the world. The goal of such analysis in visual perception is to infer the identities, forms, and spatial arrangement of objects in the three-dimensional (3-D) scene based on our two-dimensional (2-D) retinal images. *Computational approaches* to perception seek to elucidate the theoretical principles and to model the mechanisms underlying these analyses and inferential processes.

The theoretician David Marr suggested that computational accounts of perception should provide explanations at three different levels: (1) computational theory, (2) representation and algorithms, and (3) implementation. Accounts at the computational theory level clarify the purposes or goals of the computations underlying a perceptual phenomenon and explain the logic of the proposed strategies for achieving those goals. Next, representation and algorithm level accounts describe how information is represented and how it is transformed into the desired output given the input. The implementation level details the physical implementation of the transformation in neural circuits or in computers: notably, each implementation medium is associated with a set of unique constraints—for example, computers can manipulate numbers with high precision but in a step-by-step serial fashion, whereas neurons individually represent information in low resolution and slowly but process information in a massive parallel fashion. These three levels of computational account are coupled only loosely. There is wide choice available at each level, and the explication of each level involves issues that are rather independent of the other two. This entry describes stages and modules of perceptual computation, varieties of computational approaches, and contributions of computational approaches.

Stages and Modules of Perceptual Computation

Perceptual computation in our visual system can be roughly divided into three stages: early vision, mid-level vision, and high-level vision. Early vision involves the extraction of elementary visual features such as edges, color, and optical flow, among others, as well as the grouping of these features into useful aggregates. Mid-level vision deals with the inference of visible surfaces or Marr's so-called 2.5-D sketch. The objective here is to infer the geometrical shapes of the visible surfaces and the occlusion relationships between them in a visual scene. High-level vision concerns reasoning about objects, their identities, structures, and locations. It also concerns global scene information such as 3-D spatial layout and illumination direction. Although the three-stage division is motivated mainly by functional considerations, these computational stages correspond roughly to the different visual areas along the ventral visual pathway in the hierarchical visual cortex. Marr suggested that visual computation proceeds in a bottom-up, feedforward fashion, employing a series of loosely coupled, decomposable computational modules, each of which can be studied mathematically and computationally in isolation. Recent research, however, has begun to emphasize the functional roles of recurrent interaction among the different computational processes during perceptual inference.

Varieties of Computational Approaches

There are three major computational approaches in the study of perception: the *inverse optics* approach, the *dynamical system/neural network* approach, and the *statistical inference* approach. All three approaches consider perception fundamentally as a problem of *inference,* as proposed by Hermann von Helmholtz, for filling in the missing logical gap between the retinal images and the perceptual knowledge to be derived from them. Current research formulates perceptual inference in the Bayesian framework, which emphasizes the integration of prior knowledge in the inference process. To illustrate Bayesian inference, let us consider the following example. Suppose I saw a woman wandering around in my yard wearing a hat one evening when I got home. Normally, I would have concluded that it was my wife because she was the only woman in the house. On the other hand, grandma told me she would be visiting either that day or the next day. Thus, it could also be granny. In Bayesian terms, the prior probability of the woman being my wife was $p(wife) = 2/3$, and that of granny was $p(granny) = 1/3$. Now, from experience, I also knew granny loved wearing hats but my wife was not fond of it. Thus, the likelihood of observing granny in a hat was $p(hat|granny) = 0.2$, but that of observing my wife in a hat was much lower at $p(hat|wife) = 0.05$. The probability of a certain interpretation after combining the likelihood of a certain observation and the prior probability of that interpretation, called the posterior probability (i.e., $p(granny|hat)$, $p(wife|hat)$), can be obtained by the Bayes' rule,

$$\begin{aligned} p(granny|hat) &= p(hat|granny)p(granny)|p(hat) \\ &= 1/3 \times (0.2)/0.1 = 2/3 \end{aligned}$$

$$p(wife|hat) = p(hat|wife)p(wife)|p(hat) = 1/3$$

where

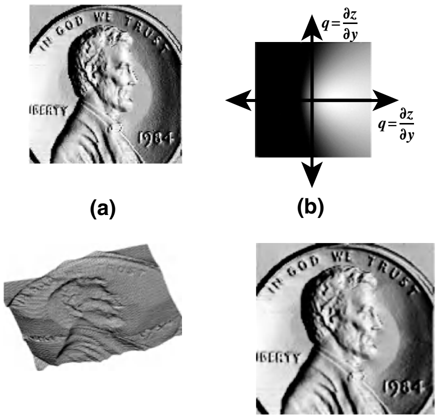
$$\begin{aligned} p(hat) &= p(hat|wife)p(wife)p(hat|granny)p(granny) \\ &= 2/3 \times (0.05) + 1/3 \times (0.2) = 0.1. \end{aligned}$$

The optimal Bayes' decision is one that chooses the interpretation with the maximum posterior probability, or the so-called maximum *a posteriori* estimate, which minimizes the probability of making a mistake. In this example, the optimal Bayesian estimate is "granny." Helmholtz suggested that such inferences are constantly taking place unconsciously in our brain, yielding perceptual conclusions about our environment.

The Inverse Optics Approach

The inverse optics approach starts the inquiry at the computational theory level by asking what physical processes are responsible for generating the observed images. This approach begins with a deterministic forward model of image formation that describes the behavior of light as it travels through space and interacts with surfaces, and then attempts to invert that model. This approach has been important in studies involving the computation of the 3-D surface geometry of objects such as depth from binocular stereo and shape from shading.

Figure 1(a) shows an image of a penny. We can effortlessly perceive the 3-D surface geometry of Lincoln's face based simply on the shading information. The inverse optics approach looks for a set of constraints based on our knowledge or models of image formation. One such constraint is that the relative luminance perceived at a particular location should depend on how the surface is oriented (slanted) away from the viewer, the lighting direction, and the reflecting properties of the surface. Figure 1(b) depicts the relationship between measured luminance and surface orientation, assuming light comes from the right and the reflectance property of the surface is Lambertian, which means any incident light will be reflected diffusely and equally in all directions. Each location in the so-called reflectance map in Figure 1(b) corresponds to a particular surface orientation. The drawn icons at the end of the x- and y-axes illustrate how the surfaces are oriented along those axes (p, q). The luminance at each point is the luminance seen at each corresponding surface orientation given a single specific lighting direction. A number of surface orientations can give rise to each particular luminance, so the surface orientation at each location in the image cannot be uniquely determined based on the measured luminance alone. Such ambiguity can be resolved with additional



(a) (b)

(c) (d)

Figure 1 The Inverse Optics Approach

Notes: (a) An input image. (b) The Lambertian reflectance map with specified lighting (from the right) that relates surface orientation (*p,q*) and observed luminance. (c) The recovered 3-D shape based on the image by the inverse optics approach. (d) The image rendered based on the recovered shape and the assumed lighting (courtesy of Brian Potetz).

assumptions such as that surfaces of objects tend to be smooth.

Optimization algorithms are typically used to find surface orientation, as defined by (*p, q*), at each image location that can explain the luminance measurement at that location based on information in the reflectance map, while requiring the estimate surface to be smooth. Figure 1(c) illustrates a recovered 3-D shape from the penny image, which is computed based on the maximum *a posteriori* estimate of surface orientation at each point in the image. Figure 1(d) shows a rendered image of the recovered shape, which almost completely explains the input image. Yair Weiss and colleagues also showed that with the prior assumptions that objects tend to move slowly, that each object is rigid and its different parts tend to move at the same velocity, the maximum *a posteriori* estimate of motion stimuli matches well to our percept, suggesting that our perceptual system is performing Bayesian inference with our percept being an optimal Bayesian estimate.

Studies based on this approach have yielded important insights into the physical constraints of the world that help shape our perception. A difficulty with this approach is that many physical processes responsible for generating images in natural scenes remain unknown or difficult to model mathematically. Furthermore, each observed image is generated by many processes (such as surface materials, 3-D surface geometry, lighting conditions) acting in concert and can be produced by a number of different combinations of these causes. Similar objects can give rise to many different images, whereas many different objects can give rise to similar images. Thus, the inverse optic approach tends to be highly under-constrained, and works only in restrictive and idealized situations.

The Neural Network/Dynamical System Approach

The neural network or dynamical system approaches focus on considerations at the neural implementation and mechanism level. Typically, neural circuits are constructed from simplified models of neurons with excitatory and inhibitory connections to explain perceptual or neurophysiological observations. The neural circuit developed by Stephen Grossberg and colleagues to explain the perception of neon-color spreading (color insert, Figure 31a) provides a beautiful example. In their boundary contour/feature contour model, the boundary detection system detects the abrupt black to red transitions in the lines as evidence of occlusion boundary. A boundary linking mechanism based on Gestalt laws completes the boundary, which in this case forms an invisible circular fence to constrain the diffusion of color from the red cross, producing an elusive reddish tinted surface percept. The intuition implemented here is that *a priori* our visual system assumes color to be constant over a continuous surface (color insert, Figure 31b and c).

This class of models has a strong intuitive and conceptual appeal to biologists and psychologists because of its obvious connections to perceptual observations, physiological data, and plausible neural circuits. The design of these models tends to be bottom-up, that is, wiring up the circuits by hand or by some learning rules, rather than being motivated "top-down" by computational theory and constraints. Hence, these models tend to offer explanations at the mechanistic levels, without necessarily explaining the underlying computational principles and rationales at the theoretical level. From a computational perspective, however, neural network model approaches do have an advantage:

Mechanistic learning rules implementing supervised and unsupervised learning can be applied to wire the circuits based on input data. That means that at least some inference mechanisms can be learned by implicitly extracting statistical regularities from the input data in the learning process without an explicit understanding of how images are formed in the natural environment. Kunihiko Fukushima's neocognitron model and Tomaso Poggio's HMAX model are well-known examples that combine unsupervised or supervised learning to achieve recognition of specific categories of objects in uncluttered scenes. Nevertheless, despite the success of such approaches in the detection of some specific object classes such as letters or faces, learning a general perceptual inference system that can fully interpret images from 3-D natural environments remains intractable with such approach because the parameter space for associating images with their underlying causes is extremely large.

The Statistical Learning and Inference Approach

The emphasis on understanding the causes of image formation in the inverse optics approach and the emphasis on extracting statistical patterns from natural scene data in the neural network approach have recently been united under the framework of statistical learning and inference with generative graphical models. The statistical approach seeks to model explicitly the statistical structures in the actual joint probability distributions between images and their underlying causes to incorporate our knowledge about the conditional independence between the causes. Statistical learning models the uncertainty about the image's underlying causes, which is easier than fitting precise physical models based on image formation.

An early success story of this approach is the modeling of the emergence of the primary visual cortex (V1) simple cells' receptive fields as being independent image feature factors for representing (or synthesizing) natural images efficiently. This idea of efficient coding, originally proposed by Horace Barlow, might be generalized to conceptualize perception as a process for seeking and learning the simplest and most parsimonious descriptions of the observed scenes. Decomposing image causes as independent factors allows irrelevant factors or causes to be ignored in efficient scene descriptions. The theoretician David Mumford suggested that the visual system computes and represents a hierarchy of efficient codes and that the recurrent feedback connections between the different visual areas provide a means for the higher visual areas to "hallucinate" interpretations that can help select and disambiguate representations in the lower visual areas. The generation of these hallucinations is not deterministic but probabilistic in nature. Each hallucination is effectively a sample drawn from the joint statistical distribution represented in this hierarchy. These distributions are encoded in the horizontal and feedback connections, furnishing *a statistical grammar* that allows the visual system to parse and interpret images in a way that relates how linguistic grammar is used to parse spoken or written sentences.

In this framework, visual processing involves iterative top-down and bottom-up processing. The broad interpretations of overall scenes (e.g., outdoor vs. indoor) are often much easier to compute than are early visual processes such as image segmentation, figure-ground segregation, and 3-D surface shape inference, so these interpretations could provide useful contextual information to disambiguate early visual processing. Top-down hallucinations based on these interpretations can exert two effects: enhancing relevant visual cues over irrelevant information or noise in the early representations, while explaining away or suppressing the redundant information in the low-level representation. Through this explaining away, the activation of the higher areas results in suppression of the overall activities in the early visual areas, resulting in a more efficient—in terms of overall neural activities in the brain—neural description of the environment. From this perspective, subjective perception corresponds to a synthesized mental representation that provides the most efficient scene description.

Learning such a hierarchical coding or statistical grammar structure is difficult. It has become more tractable when clever mathematical structures are introduced for the variables and their relationship and large amounts of data. Figure 2 illustrates an example of these ideas from the works of Stuart Geman, Song-Chun Zhu, and others. This example concerns how the concept of a bicycle could be represented hierarchically in terms of its parts and how statistical distributions in this structured hierarchical representation can be learned and used in the perceptual parsing of a bicycle. A typical bicycle is composed of four characteristic parts: seat, handle, wheels, and frame. The dependency relationship between the parts and the bicycle concept is specified by the so called AND edges, meaning that the concept of a bicycle requires the simultaneous observation of a seat AND a handle AND wheels AND a frame. The spatial relationships between these parts themselves are constrained by horizontal edges, each modeled by statistical distributions. For example, the distance between the center of the frame and the wheels follow a Gaussian distribution estimated from training data. Each part itself can have many possible instantiations— a wheel could be a thick wheel of a mountain bike OR a thin wheel of a racing bike. This relationship is expressed as OR edges in the graph, with each edge encoding the probability of a possible realization associated with that part. This AND/OR structure repeats itself iteratively down the hierarchy until the parts are basic image primitives such as corners, junctions, and edges represented in early visual cortex. Perceiving a bicycle in an image means linking image features to subparts, and ultimately to the concept of a bicycle, in the image processing process. The bottom row of Figure 2 shows the retinal input image and the ambiguous edge signals from bottom-up edge detection (left). The Gestalt laws embedded in the horizontal connections in the early areas suggest the presence of two circles, which lead to the proposal that a bicycle might be seen. This triggers the system to generate top-down hallucinations (sampling from the statistical distributions) consistent with the input. The seat and handlebar are hard to see in this particular input image, so the top-down hallucination can even hypothesize what the presence of those missing parts would look like. Such hallucinations or hypotheses can be realized as explicit mental images in the early visual cortex (three examples on the right of the bottom row). It

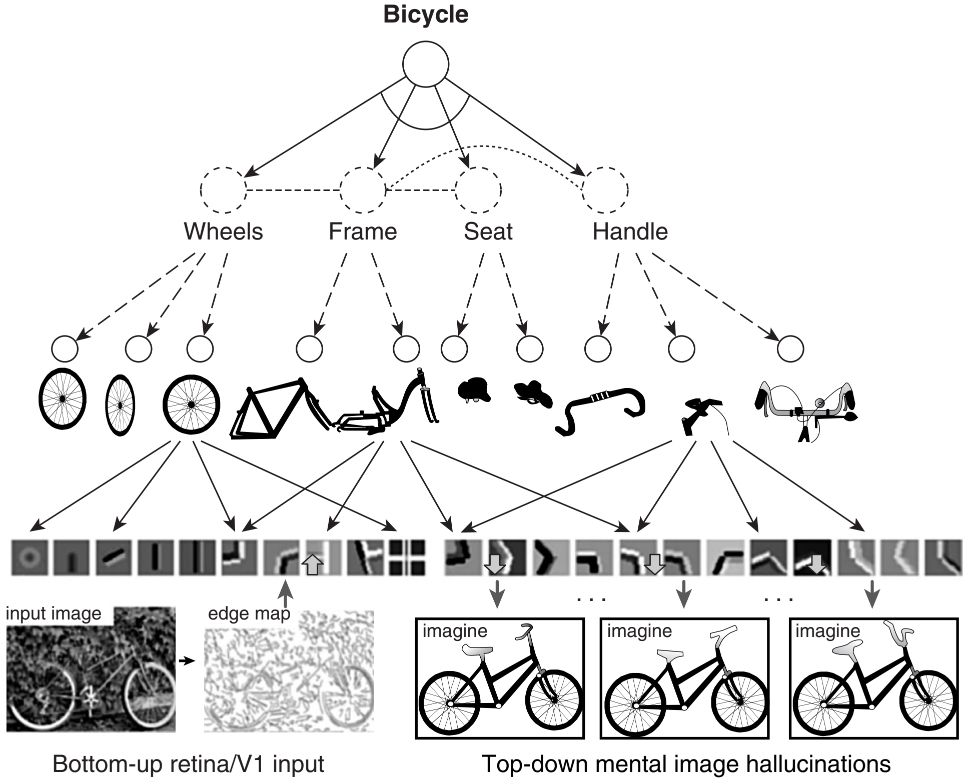


Figure 2 A Schematic of the Generative AND/OR Graph for Representing the Concept of a Bicycle

Notes: Based on Z arrows and horizontal arrows specify probability distributions relating the different structure variables. Random samples from the statistical distributions in this graph produced reasonable bicycles in different views and forms. Recognition of a bicycle in a cluttered and ambiguous natural scene (left, bottom row) is made possible by top-down and bottom-up interaction.

has been shown that bicycles in extremely ambiguous natural scenes can be segmented and recognized using this computational architecture, a feat unparalleled by other existing computational systems based simply on feedforward connections. Recently, it has been shown that this type of hierarchical network with many layers of representations can be learned in an unsupervised manner. However, it remains to be shown empirically that perceptual inference in our visual system is indeed based on these computational principles and mechanisms.

Contributions

Computational approaches have, over the years, provided a deeper, clearer, and more precise understanding of the computational principles and mechanisms underlying perception—why we perceive the world the way we do—by considering subjective perceptual phenomena as solutions to computational problems and by revealing how physical and statistical constraints on natural scenes shape our perceptual systems. Computational research has made clear the difficulties confronted by perceptual computation and have stretched the envelope of our imagination for the possible principles and mechanisms needed to resolve these difficulties. An emerging theme is that the perceptual computation can be conceptualized as a process of Bayesian inference of the underlying causes of the images. The objective of the inference is to provide us with a parsimonious and efficient description of the world around us so that we can ignore details irrelevant to the task and pay attention to details that are critical to our survival and task success. This inference could be facilitated by the mechanism of *analysis by synthesis,* which dynamically constructs the outcome of the inference as subjectively perceived mental images in our brain.

Tai Sing Lee

See also Bayesian Approach; Statistical Learning; Theoretical Approaches; Visual Scene Perception; Visual Scene Statistics

Further Readings

Geman, S. (2006). Invariance and selectivity in the ventral visual pathway. *Journal of Physiology—Paris, 100,* 212–224.

Grossberg, S., & Ennio Mingolla, E. (1985). Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. *Psychological Reviews, 92*(2), 173–211.

Marr, D. (1982). *Vision.* New York: W. H. Freeman.

Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biological Cybernetics, 66,* 241–251.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2,* 1019–1025.

Weiss, Y., Erro, P., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5*(6), 598–604.

Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences, 10*(7), 301–308.

Zhu, S. C., & Mumford, D. (2006). A stochastic grammar of images. *Foundations and Trends in Computer Graphics and Vision,* 2(4), 259–362.

COMPUTER CONSCIOUSNESS

"I'm afraid, Dave." In Stanley Kubrick's classic film *2001: A Space Odyssey,* the computer HAL uttered these provocative words as it was being shut down. They remain provocative to this day: Can a computer really experience emotions such as fear? Can it really have an *I* that is afraid to be shut down? An *I* that is the subject of genuine conscious sensations? Or must a computer, no matter how sophisticated its program and convincing its behavior, be forever devoid of conscious experiences? These are key questions about computer consciousness, questions still debated by experts and explored in blockbuster films. This entry discusses some of these questions.

The Brain as Computer

Some experts answer, "Of course a computer can be conscious. The human brain, for instance, is a computer, and it has conscious experiences. So computer consciousness is not just possible, it is commonplace."

These experts differ, however, on why, exactly, the brain can be conscious. Some are *biological naturalists,* who claim that special properties of brain biology are critical. Precisely what these properties are, and how they can generate, or be, conscious experiences, is an open question with no scientific theories yet on offer. But one implication of biological naturalism is clear: If biology is necessary, somehow, for consciousness, then any complex system that lacks biology must also lack consciousness. Because the brain is a biological computer, it can be conscious. But a nonbiological computer, such as HAL, could not be conscious, no matter how compelling its utterances.

Other experts are *functionalists,* who claim that the critical properties are not fundamentally biological, but functional. The brain can be understood as running complex programs serial, parallel, and even quantum. Certain properties of these programs are critical for consciousness. Again, no scientific theory yet explains, precisely, what these functional properties are and how they generate consciousness; perhaps concepts from information theory or complexity theory will be useful. But functionalism is clear that biology, per se, is not essential to consciousness. A nonbiological computer, such as HAL, could be conscious if it is properly programmed.

Biological naturalists assert that progress in neuroscience is required to make progress in understanding consciousness. Functionalists can agree that progress in neuroscience is important because careful study of brain function might illuminate the functional properties that are critical to consciousness. Thus, both can profitably learn from neuroscience. But they debate about how this knowledge can be used. Functionalists claim that we can, in principle, use it to build conscious, nonbiological machines. Biological naturalists disagree.

Reductive and Nonreductive Functionalism

Functionalism is by far the more prevalent view among experts today. There are many versions of functionalism, and technical nuances within these versions. But functionalists can be grouped into two broad classes.

Reductive functionalists claim that mental states are identical to certain functional states: The conditions that define the different types of mental states of a system, whether biological or not, refer only to relations between inputs to the system, outputs from the system, and other mental states of the system. The relations among inputs, outputs, and mental states are typically taken to be causal relations. However, the reductive functionalist does not claim that these causal relations *cause* mental states. Instead, this functionalist claims that mental states *are* certain functional states. In particular, states of consciousness are mental states and are thus, according to the reductive functionalist, identical to certain functional states. If a computer, such as HAL, happens to have the right functional states, then it ipso facto has conscious experiences.

Nonreductive functionalists claim that mental states arise from functional organization but are not functional states. Consciousness, in particular, is determined by functional organization, but it is not identical to, or reducible to, functional organization. Nonreductive functionalism is, in one sense, a weaker claim than is reductive functionalism because it claims only that functional organization determines mental states, but drops the stronger claim that mental states are identical to functional states. But in another sense nonreductive functionalism is a stronger, and puzzling, claim: Mental states, and conscious experiences in particular, are something other than functional states, and therefore have properties beyond those of functional states. This proposed dualism of properties raises the unsolved puzzle of precisely what these new properties are and how they are related to functional properties. However, the nonreductive functionalist does agree with the reductive functionalist that if a computer, such as HAL, has the right functional organization, then it will have conscious experiences.

Spectrum Inversion

Reductive functionalism, although controversial, is the dominant view among experts today. One thought experiment at the center of the controversy is the so-called spectrum inversion problem, which goes back at least to John Locke. He asked, in his 1690 *Essay Concerning Human Understanding,* if it were possible that "the idea that a violet produced in one man's mind by his eyes were the same that a marigold produced in another man's, and vice versa." Are the colors you see the same as the colors I see? More specifically, suppose that you and I are functionally identical.

Would it still be possible that our color experiences differ, so that, for instance, the color I experience when viewing a ripe tomato is the color you experience when viewing fresh grass?

Functionalists, both reductive and nonreductive, answer that it is not possible for two people to be functionally identical and yet to differ in their color experiences. The reason, according to functionalism, is that every mental state, and therefore every color experience, is determined by functional organization. So, if two people have the same functional organization they must have the same mental states, and therefore the same color experiences.

If it could be shown that spectrum inversion were possible, this would falsify functionalism. It would challenge whether computer consciousness is possible because most arguments in favor of computer consciousness are based on functionalist assumptions. Thus, the possibility of spectrum inversion is still widely debated.

If reductive functionalism were true, then it would in principle be possible to build a nonbiological computer, a variant of HAL, that is functionally identical to you. In this case, if you and the computer were shown the same visual scene, then the conscious color experiences of this computer, indeed all its conscious experiences, would be identical to yours.

Technology will likely evolve to the point where computers behave substantially like intelligent, conscious agents. The question of computer consciousness is whether such sophisticated computers really are conscious, or are just going through the motions. The answer will be illuminating not just for the nature of computers but also for human nature.

Donald D. Hoffman

See also Consciousness; Inverted Spectrum; Neuropsychology of Perception; Vision

Further Readings

Blackmore, S. (2004). *Consciousness: An introduction.* Oxford, UK: Oxford University Press.

Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory.* New York: Oxford University Press.

Hoffman, D. (2006). The scrambling theorem: A simple proof of the logical possibility of spectrum inversion. *Consciousness and Cognition, 15,* 31–45.

Hofstadter, D. (2007). *I am a strange loop.* New York: Basic Books.

Pinker, S. (1999). *How the mind works.* New York: W. W. Norton.

COMPUTER-GENERATED SPEECH, PERCEPTION OF

The ability to generate synthetic speech under computer control provides an important tool for the study of speech perception as well as an important technology for interacting with computers. We generally think of speech as a uniquely human ability, even if qualified by the vocal mimicry abilities of other species such as parrots and mynah birds, but computer models of speech production make it possible to generate speech signals with specific acoustic characteristics. Speech is an extremely complex acoustic signal mixing periodic and aperiodic sounds and consisting of patterns of frequency change over time, bursts of noise, silent gaps, and brief steady state patterns. Understanding how listeners quickly and effectively understand this complex signal is the primary goal of speech perception research, and synthetic speech allows systematic control of these acoustic properties. This entry describes the scientific importance of synthetic speech, models of speech production, text-to-speech synthesis, resynthesis of speech, and applications and limitations of computer-generated speech.

Scientific Importance of Synthetic Speech

Although psychoacoustic research defines and manipulates the acoustic properties of stimuli exactly using formal mathematical descriptions, this has not generally been possible for speech research. No simple mathematical description of speech can be used to characterize the sound patterns that affect listeners' perception of speech. Phonetic research seeks to identify the acoustic patterns or movements and positions of parts of the speaker's mouth that determine perception of consonant and vowel sounds (i.e., phonemes). The development of the speech spectrograph allowed researchers to measure how the sound patterns of speech change over time. The x-axis of the spectrogram (printed by a spectrograph) shows time as an utterance unfolds and the y-axis displays the frequency with the amount of energy at each time point and frequency displayed by darkening the point. Thus, dark and light visual patterns depict the different frequencies and noises in the acoustic patterns of speech. Rising pitches are displayed as lines that slant upward and falling pitches are displayed as lines that slant downward. The development of the Pattern Playback Machine at Haskins Laboratories provided a way of turning those visual patterns into acoustic speech sounds. Different patterns painted onto acetate were converted into speech containing only those acoustic properties hypothesized to change perception of one consonant or vowel into another. By making slight variations on those visual patterns, small systematic changes in speech signals could be generated where it would be difficult or even impossible for a human talker to speak with such precise control.

Models of Speech Production

Speech synthesizers are the modern computational version of the Pattern Playback. Instead of drawing acoustic patterns in visual form, speech synthesizers take as input a description of the speech signal—for example, as numerical descriptions of acoustic properties or moment-by-moment physical positions of speech articulators such as the tongue. These descriptions are then input to a computational model of speech production that generates the actual acoustic output. Some synthesizers model the movements of the mouth whereas others model the acoustic properties of speech production through filters and resonators.

Despite the differences between the underlying models, however, speech synthesizers are used less often to test those models than to parametrically control the acoustic properties of speech tested in perceptual experiments. Synthesizers permit precise control of the pattern properties of speech, synthetic speech has made possible studies how these properties are used in recognizing spoken consonants, vowels, and words, how this ability develops from infancy, and how human perception of speech differs from perception of speech signals by nonhuman animals. For example, it is possible to make a series of test stimuli varying from one vowel to another (e.g., from EE as in *beet* to IH as in *bit*) or from one consonant to another (e.g., from B as in *bit* to P as in *pit*). In the case of vowel stimuli, this can be accomplished by varying the duration of the vowel or by changing the frequency of one of the components of the vowel. For consonant differences, it is possible to vary the timing relationships among different acoustic properties such as a burst of noise and the vowel of the syllable, or by changing frequency properties. In this way, small acoustic changes can be produced that can be classified by listeners (adult humans, infants, chinchillas) to understand the relationship between the acoustic properties and perception.

Text-to-Speech Synthesis

Speech synthesizers that model speech production can also be combined with linguistic rules to produce text-to-speech systems. A text-to-speech system (TTS) takes unrestricted text such as this article and "reads" the text aloud, or at least renders the text into a speech waveform. Linguistic rules are used to translate the text input into a symbolic representation of consonants and vowels, and then to translate these phonemes (organized into words and sentences) into acoustic parameters that would drive the speech synthesizer to produce acoustic output. In this respect, a TTS is more like a reader than like an extemporaneous speaker. Although speech synthesizers that work at the level of modeling speech production are important for fine systematic control of the acoustic properties of consonants and vowels, text-to-speech systems are more useful to study the perception of spoken words and sentences.

Resynthesis of Speech

Speech can also be synthesized directly from the model of a natural utterance such as a sentence. Any utterance can be analyzed (using a modeling technique called linear predictive coding) to measure acoustic properties of the talker's voice and the moment-by-moment acoustic properties of the talker's vocal tract. These measurements can then be input to a model that generates an acoustic speech signal that is closely matched to the original speech. These parameters can then be modified to produce variations in consonants or vowels or speaking rate or intonation, or even whether the talker sounds male or female. This approach generates intelligible and natural-sounding speech that can be used to study the perception of emotion and speaker attitudes or aspects of narration.

Applications and Limitations

In general, synthetic speech can sound artificial. Listeners detect that such speech is not naturally produced, either by the presence of extraneous sounds or by funny intonation or unintelligible syllables. This has not limited the usefulness of synthetic speech for the study of speech perception. Moreover, the intelligibility of synthetic speech has reached the point that it is routinely used in a variety of commercial applications including voice response systems in automated teller machines, announcement systems in public transportation, vocal prosthetic devices for people who cannot talk, and reading machines for the blind.

Howard C. Nusbaum

See also Human–Machine Interface; Language; Speech Perception; Statistical Learning

Further Readings

Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2003). Consolidation during sleep of perceptual learning of spoken language. *Nature, 425,* 614–616.

Francis, A. L., Baldwin, K., & Nusbaum, H. C. (2000). Learning to listen: The effects of training on attention to acoustic cues. *Perception & Psychophysics, 62,* 1668–1680.

Francis, A. L., Fenn, K., & Nusbaum, H. C. (2007). Effects of training on the acoustic-phonetic representation of synthetic speech. *Journal of Speech, Language, and Hearing Research, 50,* 1445–1465.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74,* 431–461.

Taylor, P. (2009). *Text-to-speech synthesis.* Cambridge, UK: Cambridge University Press.

COMPUTER GRAPHICS AND PERCEPTION

Computer graphics has become pervasive in visual media and is the core technology in video games, whose revenues are now higher than ticket sales for movies. Movies wholly generated using computer graphics appear with regularity, and movies commonly have backgrounds or characters created using computer graphics. Print and broadcast media have been similarly affected, and computer graphics virtual worlds are becoming commonplace on the World Wide Web. Because the product of computer graphics systems is intended to be viewed by people, it is important to understand computer graphics in a perceptual context. Although perception of computer graphics is similar to that of photography, film, and the graphical arts, important aspects of image creation and display are unique to computer graphics and must be understood to understand the perceptual effectiveness of computer graphics. This entry describes how computer graphics is generated, limitations of display devices, spatial information in computer graphics imagery, pictorial perception, and non-photorealistic rendering.

How Computer Graphics Is Generated

A computer graphics program can be thought of as a synthetic camera that captures an image or sequence of images of a scene that is described virtually in the computer. The virtual description of the scene (model) contains information about the geometry of surfaces and objects, material properties affecting the appearance of surfaces and objects, and characterizations of the light sources. The complexity with which illumination and surfaces interact to produce patterns of light in the real world is so great that computer graphics must necessarily operate using at best rough approximations to actual physics. These approximations must strike a balance between appearance and computational efficiency.

Usually, computer graphics generates or renders images using the same perspective projection process that occurs in real cameras and in the human eye. In computer graphics, a mathematical specification of perspective projection is used to map positions in the scene to their corresponding image locations. A variety of techniques are available to ensure that objects and surfaces in the scene hidden (occluded) by nearer objects and surfaces do not affect the appearance of the final image.

In the real world, there is variation across surfaces in color and reflectivity. Real-world surfaces also have fine-scale geometric roughness. Computer graphics models surface variations in color and reflectivity using *texture maps,* which may be hand painted in an image-editing program, captured by a camera, or produced in some automated manner. Similar techniques are available to represent geometric textures in ways that can be efficiently used by rendering algorithms. These kinds of textures are widely used in both computer games and movies (see Figure 1).

The manner in which light reflects off surfaces is modeled using some form of *bidirectional reflectance distribution function* (BRDF), which relates the percentage of incident light that is reflected to the orientations of light sources and viewpoint relative to the surface. The nature of the most commonly used rendering process is such that adding shadows to the generated image is quite costly, though it is now feasible to do in real-time on high-end graphics cards. A fully accurate image must also account for indirect lighting that reflects in series from more than one surface. This is not currently possible for interactive systems such as vehicles simulators and video games, but is often done for computer graphics–generated movies.

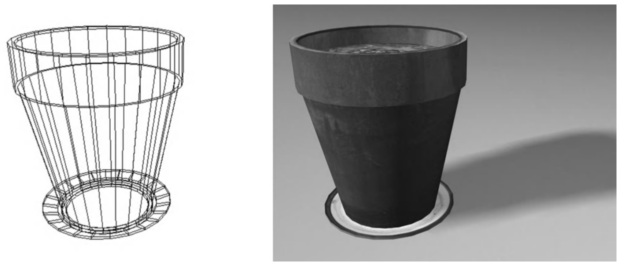


Figure 1 Example of Computer Graphics Imagery

Source: Figure courtesy of Leah Wouters.

Notes: On the left is an indication of lines making up the geometric model. The right image is produced using perspective projection, lighting, texture mapping, and shadow.

Limitations of Display Devices

Computer graphics images are usually presented for viewing using either some sort of display device such as a monitor or video projector or in printed form. This has important perceptual implications, which are shared with other visual media such as photography and movies that use the same mechanisms for presentation. Under normal viewing conditions, the human visual system is able to see detail two to three times finer than can be displayed on computer monitors or high-definition televisions (HDTVs), when these are viewed at normal distances. Dynamic range, sometimes referred to as contrast ratio, is the difference in brightness from the darkest black to the lightest white in an image. Except for a few exotic display technologies, it is not possible to present images with a dynamic range close to that encountered by the human visual system in everyday life. With adaptation, the human visual system is sensitive to a range of brightness of about eight orders of magnitude, and can easily deal with naturally occurring scenes in which brightness varies by a factor or 10,000. High-quality computer and HDTV monitors can display a range of brightnesses covering about three orders of magnitude and are incapable of simulating the light levels associated with either bright moonlight or darker, or outdoors in daylight. Print media are far more restrictive in the brightnesses that can be conveyed. Similarly, the range of colors that can be seen by the human visual system far exceeds the range of colors (color gamut) that can be displayed by computer monitors, image projectors, or print media.

To deal with the problem of presenting images that have luminances and colors not displayable on a monitor or in print media, some sort of *tone mapping* operation is required. Tone mapping involves generating an image that gives the impression of the real scene, but using only the available dynamic range of the display device. The simplest form of tone mapping—used in many computer graphics systems—is to multiplicatively scale brightness levels to bring them within the required range. Because this may involve reducing contrast by factors of 1,000 or more, the result rarely appears realistic. More sophisticated techniques exploit the insensitivity of the human visual system to show variations in brightness across the field of view by adaptively increasing or decreasing brightness and contrast in local areas of the image. Done correctly, this can dramatically increase the subjective sense of dynamic range in an image without any modifications to the display device.

Spatial Information in Computer Graphics Imagery

Vision is a powerful source of information about the three-dimensional structure of the visible environment. Much of this information is associated with what are called *pictorial cues,* which can be extracted from single, static views of the world. Pictorial information includes various perspective effects, shading, shadows, and the patterns that occur when one object partially hides another. Visual motion is pervasive in human experience, and often provides a powerful source of information about the three-dimensional structure of the world. Binocular stereo supports perception of depth, surface orientation, and occlusion, particularly at nearer distances. Focus (accommodation) can also indicate depth, but only at distances up to a meter or two.

The term *perspective* is used to apply to a range of visual effects, all associated with the fact that the size of the retinal projection of a given sized entity in the environment is smaller the farther away the entity is from the viewer. The best-known perspective effect is that parallel lines in the world, when viewed from an oblique angle, converge to a vanishing point as when looking down a long straight road. A retinal image of converging lines is often perceived as a surface receding away from the viewer. The location of the vanishing point affects the perceived slant of the surface and the perceived relative distances to various points on the surface. The world is full of *visual texture,* consisting of repeated patterns such as grass, forests, brick, carpet, and the like. When a textured surface is viewed obliquely, the patterns making up the texture become finer and finer with distance. The visual system can exploit this effect, perceiving textures that smoothly vary in scale as indications of a slanted surface.

Computer graphics renders geometric models using perspective projection as an integral part of the process. The generated image of geometric entities will correctly convey the visual cues that underlie perspective-based space perception. The effectiveness of these cues in generating a perception of depth and surface orientation will depend partly on the accuracy of the geometric models used to generate them, though the nature of perspective cues is such that they do not depend strongly on the detailed shape of objects and surfaces. Although most computer graphics systems render geometry in a perspectivally correct manner, the same is not true for texture maps used to specify the detailed appearance of surfaces. Approximations used to render texture maps efficiently slightly distort the spatial structure of the texture and can significantly deviate from the correct association of the scale of the texture with viewing distance.

As indicated previously, the brightness of a point on a visible surface seen by a viewer is a function of the brightnesses and locations of light sources, the reflectance of the surface, and the orientation of the surface with respect to the line of sight. In almost all cases, visible brightness changes substantially when the relative orientation of light source and surface is changed. For most materials, changes in the orientation of the line of sight to the surface also produce significant changes in brightness. The visual system often interprets smoothly varying changes in brightness of a viewed surface as being caused by smoothly varying changes in the orientation of the surface. These spatial changes in surface orientation are typically perceived as resulting from surface curvature. Sometimes discontinuous changes in brightness are interpreted as discontinuous changes in surface orientation, but visual interpretation of brightness discontinuities is complex and not well understood. The lighting models used by most computer graphics rendering systems approximate the change in brightness caused by changes in surface orientation, as indicated by the underlying geometric models or bump maps. The effect often results in a compelling sense of three-dimensional surface shape.

Shadows are another source of visual information about the spatial structure of the world. Most often, shadows indicate position indirectly. Objects casting shadows onto extended surfaces tend to appear at a distance from the viewer corresponding to the perceived distance of the location on the surface where the shadow falls. The way in which most computer graphics systems create images makes it extremely difficult to generate shadows in the correct location with the correct shape. Real-world shadows have a smooth transition for lit to unlit areas (penumbra), which is also difficult to render accurately. In computer graphics, this is referred to as a soft shadow. Algorithms for generating perceptually plausible approximations of shadows have been developed. These are most often used in situations such as computer-generated movies, where computation need not be interactive, but are also available on some high-end PC video cards.

Computer graphics can be used to produce animations. As in the real world, visual motion typically produces a strong sense of three-dimensionality, along with the appearance of object and viewer movement. The most direct way to animate computer graphics is to generate a sequence of image frames based on changing geometric models and viewpoints, and then display these images at a rate such that smooth motion is perceived. For movies, substantial time is often used in generating each frame because the resulting imagery is stored and then played back on video or film. Real-time interactive graphics requires that video frames be generated at roughly 30 frames per second or faster before the motion looks continuous. A few exotic display devices track the viewer's head position and use this to control the viewpoint in the graphical rendering, producing a strong sensation of motion parallax.

Although binocular stereo is a powerful three-dimensional (3-D) cue in normal vision, few display devices support viewing of computer-generated stereo images. In the human visual system, the eye motions that are part of stereo perception are affected by accommodation (focus of the eyes) to locations of interest. Even though stereo image displays can vary the views from the two different eye positions based on depth, accommodation is consistent with the fixed distance at which the images are displayed, reducing the effectiveness of the displays. Despite the limitations of stereo image displays, they are reappearing in 3-D movies.

Pictorial Perception

Although perception of computer graphics and other graphical media involves the same set of pictorial cues as does viewing the real world, there are important ways in which visual perception of pictures and visual perception differ. Motion is sometimes missing, and non-pictorial cues such as stereo and accommodation are almost always absent. All of these are important sources of spatial information. The visual extent of a viewed display screen or printed graphical image typically ranges from 20 degrees to about 45 degrees, much less than the near 180-degree horizontal field-of-view of the human visual system, though there is disagreement about the exact effect of limited field-of-view on space perception. Except for a few exotic display technologies mostly limited to vehicle simulators, there is visual information available indicating the location of the two-dimensional surface on which the image is displayed. This information conflicts with the pictorial cues in the image itself, typically resulting in compression of apparent depths in the rendered scene. Finally, perspective cues in a rendered image will be distorted if the location from which the image is viewed, relative to what is depicted in the image, differs from the location of the synthetic camera used in rendering, relative to the scene model (i.e., if the observer views the scene at an angle).

Nonphotorealistic Rendering

A relatively new type of computer graphics produces technical illustrations or artistic images that are not realistic. Such images are useful in applications that use the analogous human-produced graphics. For example, an automotive manual might use a line-based graphics rather than a real or synthetic photograph, and an animated film might use images with lines and brushstrokes. Because this type of graphics is distinguished from the more common realistic graphics, it is usually called nonphotorealistic rendering (NPR). A particular challenge in this area is that it often concerns itself with image-space features such as brushstrokes, so maintaining a good appearance under change of viewpoint or object position is a challenge.

William B. Thompson and Peter Shirley

See also Digital Imaging; Pictorial Depiction and Perception; Spatial Layout Perception, Psychophysical; Virtual Reality: Vision

Further Readings

Gooch, B., & Gooch A. (2001). *Non-photorealistic rendering*. Wellesley, MA: A. K. Peters.

Mortenson, M. (2006). *Geometric modeling* (3rd ed.). New York: Industrial Press.

Rogers, S. (1995). Perceiving pictorial space. In W. Epstein & S. Rogers (Eds.), *Perception of space and motion* (Vol. 5, pp. 119–163). San Diego, CA: Academic Press.

Shirley, P., Ashikhmin, M., Gleicher, M. Marschner, S. R., Reinhard, E., Sung, K., et al. (2005). *Fundamentals of computer graphics.* Wellesley, MA: A. K. Peters.

COMPUTER SPEECH PERCEPTION

Many science fiction writers and electrical engineers have envisioned that one day people will be able to talk to appliances. “Washing machine on” would be the vocal shot to be heard round the world liberating humanity from mechanical enslavement to buttons and knobs. However, although the current performance of speech recognition systems might be compared favorably to the effectiveness of commanding a dog to sit—a significant improvement from past decades when recognition performance was more like commanding a cat—talking to appliances is not the killer application that is desired by most people. But even at the current level of performance, there are a number of important successful applications for computer speech recognition systems. Moreover, understanding why computer speech perception has not reached the level of human performance can give insight into the basic theoretical problems that still need to be solved to understand human speech perception.

In principle, by using speech technology, every telephone could become a computer system. The phone could allow access to information stored in databases, could allow the entry of data into such databases, and could be used for a number of applications that take place using a keyboard, a mouse, and a visual display screen. For every telephone to become a net-connected computer terminal, we simply need to replace the keyboard and mouse with speech input through a recognition system and voice output using synthetic speech. However, the level of speech recognition performance is still well below the requirement of speaking to a computer using natural speech as if talking to a human agent. This entry discusses the problem of speech recognition, speech understanding, using speech recognitions systems, the human factors problem, and learning about speech perception from speech recognition.

The Problem of Speech Recognition

Human listeners recognize fluent sentences under adverse listening conditions with a high degree of accuracy. In the presence of background noise varying from –6 decibels (dB) SNR to +6 dB signal-to-noise ratio (SNR), we can recognize spoken words in the range of 60 to 90% correctly, with fluent sentence intelligibility even higher. Computer speech recognition systems cannot operate at this level of performance under comparable circumstances with a similar high degree of uncertainty about what might be said by a talker. One reason for this is the knowledge we bring to the problem of speech recognition. Human listeners use knowledge of their phonology (sound patterns and the organization of sound patterns), the lexicon (word knowledge), and syntax to aid in recognizing spoken words. We use background knowledge about the domain of discourse to constrain interpretations. Moreover, perhaps, regarding our ability to recognize speech even in isolated syllables, we can use our substantial experience producing and perceiving speech to aid recognition. The acoustic patterns of speech sounds are restructured at different speaking rates, across talkers’ vocal tract architecture and articulatory control, and in different phonetic and linguistic contexts. A single acoustic speech pattern may be ambiguous between different phonetic categories, and any single phonetic category may have multiple acoustic realizations in production. This many-to-many mapping between speech sounds and linguistic categories is a fundamental problem that humans solve with linguistic knowledge, cognitive processing, and experience with speech.

Speech recognition systems before the 1970s stored a fixed and limited set of utterances that could be recognized (e.g., a vocabulary of commands or digits or letters) as a collection of acoustic properties. A vocabulary was defined when each utterance was recorded by a talker into the system and was then reduced to a set of acoustic features such as spectral properties at different times. These features were represented as a template for the utterance that each spoken input was compared with. The spoken input was analyzed acoustically in the same way as the stored representations and the specific spectral features were compared over the duration of the utterance to derive a similarity score. The most similar template would be chosen as the candidate for the recognition of the input. For vocabularies such as digits where the words differed acoustically, and for isolated words spoken carefully, such systems could achieve 95% correct recognition for a single talker whose speech provided the templates, when there was no noise in the environment. However, deviations from these narrow operating constraints could reduce performance substantially so that users became frustrated easily at the repeated failures of recognition. For example, using letters instead of digits includes the items B, C, D, E, G, P, T, V, and Z, which are easily confused, and recognition performance could easily drop to 20% correct. Having a head cold or speaking quickly could reduce performance significantly. Performance of a system would generally be measured as percent correct recognition or percent error—reporting percent error in the speech engineering community has always been popular because small reductions in error rate appear larger than do small increases in accuracy once performance is over 80 to 90% correct. Recognition errors could be scored as substitution errors in which the wrong response was given or as recognition failures. Because these speech recognition systems depended on a comparison of the acoustic properties of stored templates of a vocabulary to each input utterance, performance was fragile and different from the robust performance of human listeners.

All speech recognition systems must cope with the problem of variability in the relationship between acoustic patterns and linguistic categories. Most speaker-independent recognition systems have vocabularies that are determined by the system designer and the acoustic representations are compiled into the system from a large database of speech that is analyzed into spectrotemporal features and stored as part of the system. The vocabulary can be designed to maximize linguistic contrast among the items (still satisfying the needs of the intended task), and the acoustics can be measured and representations chosen that for a large database of talkers maximizes discriminability. For simple discriminable vocabularies such as digits, these systems can easily achieve levels of performance over a telephone that is near human recognition—more than 95% correct. By contrast, speaker-dependent recognition systems allow the user to define a vocabulary and record the speech tokens that are analyzed to establish the vocabulary. This process can be iterative, in which the system prompts for specific vocabulary items, tests recognition, and then recalibrates by prompting for more training tokens. In this way, some systems have been developed as hybrids starting with a large database of speech for a large vocabulary but, upon determining through user interaction that substitution errors are too frequent (e.g., more than 15%), can prompt the user to produce some training speech that can be used to tune or adapt the initially speaker-independent vocabulary to function more accurately with a particular user.

Speech Understanding

In 1971, the Advanced Research Projects Agency (ARPA, now the Defense Advanced Research Projects Agency or DARPA) initiated a research competition to develop speech understanding systems. The emphasis in this project reflected a shift from previous engineering attempts at speech recognition—matching acoustic speech patterns to stored representations of utterances—to speech understanding. Speech understanding was defined as successful transmission of an intended message from talker to computer, even if recognition of the words or phonemes was poor. The shift away from pattern recognition of speech sounds to understanding of intended communication reflected two important trends in speech engineering. First, it was recognized that the linguistic content of an utterance was not completely represented by the phonetic transcription of that utterance alone. Phonemes could be omitted, reduced, or misarticulated, syllables or words dropped, and human listeners can still understand speech. Second, this entailed a shift in technical approach from pure signal processing and pattern classification methods to new developments in artificial intelligence and the successes of expert systems and natural language processing in the 1970s.

The goals of the 5-year competition were to develop systems that understand connected speech using a vocabulary of 1,000 words with a limited sentential syntax in a particular task domain from multiple cooperating talkers with less than 10% error. Four systems were completed by 1976 with one clear winner. Harpy could understand sentences from five different talkers (after being trained on 20 sentences from each) with a 5% error rate. The model was built to represent sentences by a network of about 15,000 states corresponding to short-term (10 milliseconds [ms]) power spectra in the sequences that would be permissible for all acceptable sentences. The network compiled all phonological, lexical, syntactic, and domain-specific constraints into a single model, turning comprehension into a statistical classification problem. The 20 training sentences were used to tune the acoustic state representations to specific talker vocal characteristics. Unlike other systems that represented expertise in words or sentences or domain knowledge as separate processing systems that interacted to determine recognition, Harpy precompiled all this knowledge into a single finite-state network of acoustic features with transition probabilities estimated for the nodes based on talker-specific acoustic properties and the linguistic knowledge of words and sentences. Thus, all the linguistic knowledge and knowledge of the domain of discourse (e.g., chess), was compiled by programmers into the system.

One important consequence of the ARPA speech understanding project was the insight that recognizing the acoustic patterns of speech did not have to be perfect or even the primary source of performance accuracy. This brings computer speech perception more into line with 50 years of research on human speech perception showing that lexical and syntactic and semantic constraints on the meaning of sentences can improve overall understanding of speech even at poor signal-to-noise ratios. We understand a sentence such as “A stitch in time saves nine” better than “The old corn cost the blood” in noisy environments, and this needs to be part of computer speech recognition systems. Furthermore, by placing the problem of speech understanding in a particular task domain (e.g., flight reservations or chess playing), knowledge about the kinds of sentences and words used in that domain, the kinds of linguistic interactions, and the meaning of those interactions relative to the goals of the speaker could all be used to make up for the limitations of acoustic pattern recognition of speech. This resulted in the development of computer systems that were much more successful at higher-level tasks than could previously be achieved. However, the addition of new constraints also introduced new kinds of errors. One apocryphal story from the competition described Harpy playing chess. When the computer’s king was completely boxed in and threatened with checkmate at the next move, the situational constraints were so strong on what the talker could logically say that Harpy would resign even if the human talker merely coughed instead of describing a move. In other words, Harpy could “hallucinate” understanding simply based on what the most expected utterance would be regardless of what was actually said. It is important for the constraints provided by expectation and knowledge not to override the actual sensory information in the speech signal.

Using Speech Recognition Systems

Speech recognition technology is needed for several general applications and has been used for with varying degrees of success: Voice interaction with computer systems over the telephone, information transfer in eyes-busy, hands-busy situations, as a prosthesis for people with manual motor impairment, dictation, and in “command and control” systems. As a general rule, if people are unable to use their hands for typing or controlling devices, speech is viewed as the logical substitute. Thus, we can think about speech as useful for transmitting information and as a way of controlling computers and machinery.

To date, arguably the greatest impact and success of speech recognition technology has been to improve the effectiveness of human telephone operators by partitioning their tasks into those that can be responded to by a computer and those that require human interaction. Reducing the need for human operators in some information transactions such as retrieving flight status or financial information has become the paradigm case of the successful application of speech recognition systems. In cases of number entry (e.g., flight numbers, bank account numbers) and simple isolated words and phrases, speech recognition can work effectively and efficiently, even over the phone with potential acoustic distortion of the speech signal and with great variability in the vocal characteristics of talkers. This kind of application generally combines aspects of information transfer and control through menu selection. These kinds of interactions allow for only a few alternative words at any point in the interaction, thereby significantly constraining the recognition problem. This has proven to be an important aspect of the successful application of speech recognition systems: Task-level constraints are important in the successful application of speech recognition systems.

The Human Factors Problem

When first conceived for real applications, speech technology was thought to be transparent as a user-system interface, at least by comparison with a mouse or keyboard, because everyone knows how to talk. However, it has become clear that every interface technology introduces its own problems that must be studied systematically. For example, speech recognition systems make errors, and there must be some way of communicating recognition uncertainty to a talker and for the talker to signal to the computer that a recognition error has been made. Recognition systems require careful articulation and some care about physical position with respect to the microphone, control of voice levels, consideration of ambient noise sources and levels, and simple experience in using these systems. Some talkers (called "sheep") are more successful in talking to speech recognition systems than are others (called "goats"), and general knowledge about interaction with technology and computers interacts with this difference in ability.

Human-to-human speech communication is generally successful and robust but not on the basis of the recognition of the speech signal alone. We can ask a speaker for clarification if we do not understand a word or sentence. Listeners can repeat part of a talker's phrase to seek confirmation from the talker. Speakers can increase redundancy by rephrasing or repeating ideas or words, and can slow speech down, emphasize important parts, and increase the clarity of production. Speakers use these modifications of speech production to increase listener comprehension secure in the pragmatic knowledge of a common set of rules for speech communication. However, if speakers engage in these natural behaviors with a computer speech recognition system, this would increase pattern variability and decrease performance. Thus, speakers fight natural tendencies and experience and must relearn the rules of conversational interaction for recognition systems.

Although these problems are seen in isolated utterance recognition systems, computer dictation systems present all these problems and more. Today it is certainly possible to dictate a document to a computer, although the computer will not match the facility of a moderately skilled typist in entering text. Dictation systems can be tuned, through a training regime (i.e., saying specific training sentences to the system when prompted), to the specific vocal characteristics of a single talker, but this cannot overcome the loss of the constraints that benefit small vocabulary recognition systems. Dictation is given as fluent sentences rather than as isolated words. Variation in pronunciation from time to time, because of distraction, fatigue, carelessness, or simple moment-to-moment motor variability in articulation results in acoustic patterns that are less stable for words that may have many similar "neighbors." The vocabulary is relatively unconstrained so that acoustically similar words may be used (e.g., think and thing, bog and dog). The syntax is relatively unconstrained so that the number of possible words that may follow any particular word is large. Speakers may pause, hesitate, restate, and revise an utterance after false starts. Thus, the set of utterances is poorly defined and highly variable in linguistic and acoustic terms. Unfortunately, it is difficult to predict the actual performance of a dictation system in the real world, and impartial systematic studies are scarce. System vendors will generally report high levels of accuracy (more than 95% correct) for cooperative users, although anecdotal user experience is much less satisfactory. Although it is difficult to carry out controlled dictation tasks as such, studies examining recognition of fluent speech have reported word recognition error rates around 4% for human listeners and 43% errors for computer recognition.

Human dictation taken by humans is also less successful by comparison with the more constrained interactions that take place with the human telephone operator tasks that speech recognition systems have been applied to. Dictation is a much more open-ended process wherein the constraints on any particular word at any point in time are provided by the structure of language and the message of the talker. As in the ARPA speech understanding project, to make use of such constraints effectively, it would be necessary to model the structure and use of phonemes, words, and sentences, to understand the talker's ways of speaking, and have some kind of understanding of the talker's topic. With human dictation, the normal rules of vocal communication can apply and the listener can request clarification and correct errors in ways that a computer cannot. Following the success of the Harpy system in the ARPA project, computer dictation systems typically model language using statistical networks called Hidden Markov Models that approximate these constraints based on a limited sequence of words. (The "hidden" part is the assumption of postulated states such as phonemes that serve as part of the structure of the statistical Markov model.) For example, all the linguistic knowledge that constrains the recognition of spoken words in dictated sentences is often modeled by the transition probabilities of sequences of three or four words. Although sequences of four words approximate locally comprehensible sentences, humans can easily distinguish sentences that only obey this constraint from natural sentences. The database of language from which these transition probabilities are estimated may be specific to a certain kind of discourse (e.g., business correspondence) and requires millions of words of correspondence to develop statistically plausible models even for such a small sequence of words. The implications of this is that, although a particular domain of discourse will represent the broad structural constraints of the language of that discourse and thus model sentences generally well, the specific statistics of word usage and phrases in the domain of the correspondence database can influence recognition performance substantially. Words that are rare in one domain and are potentially confused (e.g., mesa and peso) might lead to substitution errors in which a letter home is transformed by the recognition system to be less about scenery and more about banking. Similarly low-frequency constructions that are perfectly legal, such as, "He saw that that was there," will have recognition errors.

Learning About Speech Perception From Speech Recognition

The development of speech recognition technology has only been weakly influenced by knowledge of human speech perception. The application of linguistic knowledge and task knowledge is one clear example and the use of recognition "competition" within a "neighborhood" of similar utterance representations is another. However, there has been scant effort to apply what we know about the details of human speech perception and spoken language understanding to improving these systems. Moreover, little has been learned about human speech perception from the performance limitations of speech recognition systems. The use of statistical learning and statistical modeling in theories of speech perception may arguably be one such example. Some theories of human speech perception have been influenced by these notions, which suggest speech perception is based on these statistics rather than abstract linguistic categories. There have been attempts to improve speech recognition systems by using more psychoacoustically plausible acoustic pattern representations. However, this engineering application of perceptual research has demonstrated the need to better understand the relationship between peripheral auditory processes and subcortical and cortical neural mechanisms of speech perception. The few theories of speech perception that are implemented computationally and take real speech input could be thought of as the first steps toward being able to bridge the engineering approach and the psychological approach. Unfortunately, the performance of such systems is typically much worse than is that of commercial speech recognition systems. The era of the ARPA project and the immediate aftermath produced the greatest substantive interactions between speech engineers and basic speech researchers. These interactions led to new conceptualizations of the nature of the mental lexicon and the process of word recognition, as well as interactions between top-down and bottom-up processes in spoken language understanding.

Howard C. Nusbaum

See also Computer Vision; Human–Machine Interface; Language; Speech Perception; Statistical Learning

Further Readings

Jurafsky, D., & Martin, J. (2000). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition.* Upper Saddle River, NJ: Prentice Hall.

Klatt, D. H. (1977). Review of the ARPA speech understanding project. *Journal of the Acoustical Society of America, 62,* 1324–1366.

Manning, C., & Schutze, H. (1999). *Foundations of statistical natural language processing.* Cambridge: MIT Press.

Nusbaum, H. C., & Magnuson, J. S. (1997). Talker normalization: Phonetic constancy as a cognitive process. In K. A. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing.* New York: Academic Press.

Nusbaum, H. C., & Pisoni, D. B. (1987). Automatic measurement of speech recognition performance: A comparison of six speaker-dependent recognition devices. *Computer Speech and Language, 2,* 87–108.

Scharenborg, O. (2007). Reaching over the gap: A review of efforts to link human and automatic speech recognition research. *Speech Communication, 49,* 336–347.

Waibel, A., & Lee, K. (Eds.). (1990). *Readings in speech recognition.* San Francisco: Morgan Kaufman.

Weinschenk, S., & Barker, D. (2000). *Designing effective speech interfaces.* New York: Wiley.

COMPUTER VISION

Computer vision is a field whose goal is to take as input images or video streams and to extract useful information about the objects and surfaces in the scene that gave rise to the image or video. Humans employ vision to acquire knowledge about the external world and to guide a wide range of actions, and the grand objective of computer vision research is to provide machines with similar capabilities. These can be divided into the following:

> *Segmentation.* This is the partitioning of an image into sets of pixels that correspond to single objects or parts of objects. In a video stream, tracking is the process of matching these groups over time.
>
> *Recognition.* This involves categorization, which operates at a more generic level (This is the image of a face), as well as identification of specific instances (This is the image of Barack Obama's face).
>
> *Reconstruction.* This is the recovery of depth, surface orientation, and shapes of the objects and surfaces in the scene that gave rise to the image.
>
> *Controlling action.* These include visual guidance of locomotion, such as of a mobile robot in a building, or a car on a freeway, as well as manipulative actions such as reaching, grasping, or insertion.

This entry discusses each of these in turn and then closes with a brief overview of the past and future progress of computer vision.

Segmentation

Although the input to the human visual system is just a collection of values associated with outputs of individual photoreceptors, we perceive a number of visual groups, usually associated with objects or well-defined parts of objects. This ability is equally important for computer vision, where the term *segmentation* refers to the partitioning of the image into these groups of pixels. When extended to spatiotemporal data, the term *tracking* is often used to refer to the process of matching these groups from one time frame to the next. This has a number of applications. In monitoring people at an airport or traffic in a freeway, the basic process involves segmenting out individual people or cars and tracking them over time. Grouping can be of curves, not just regions; the process of tracing axons and dendrites in neuroscience research to infer neural network connectivity is another example. Grouping and tracking can be viewed as two aspects of the same process, that of spatiotemporal segmentation of a video stream.

Computer vision techniques for segmentation are based on algorithms that implement various grouping factors first studied comprehensively by the Gestalt school of visual perception. The single most important grouping factor is similarity—pixels that have similar brightness, color, or texture tend to be grouped together. Conversely, sharp changes in one or more of these features signal boundaries between regions. This motivates edge detection—a technique based on marking boundaries where neighboring pixels have significant differences in brightness. This has a natural generalization to color. However, if objects are textured, such as a tiger against a grassy background, then brightness or color edges may be less useful than texture edges for finding object boundaries. These are found by computing descriptors of image patches centered at different pixels and computing the locations where these descriptors change most sharply. Examples of texture descriptors might be histograms of oriented edge responses or filters modeled after the responses of cells in the primary visual cortex of primates.

Similarity is only one of the factors that can promote grouping. Curvilinear continuity is another powerful grouping factor, suggesting linking edge segments whose directions are consistent with being along a smoothly curving contour. This is an example of progressing from local measurements to global groupings, for which a variety of mechanisms have been explored in computer vision.

Currently, the most effective techniques for going from local edge measurements to global segmentation decisions are based on mathematical formulations such as Markov random fields, or normalized cut techniques based on spectral graph theory. Markov random fields provide a probabilistic formalism—the goal is to find the most probable grouping based on a model that makes it more likely for pixels to be in the same group if their features are similar, and for neighboring pixels to be more likely to be in the same group. Spectral graph theoretic techniques have a different viewpoint—pixels are nodes in a graph, and edges between pixels have affinity weights that are measures of pixel feature similarity. The goal is to find groups that have high intra-group affinity and low intergroup affinity, which is done by finding eigenvectors of a matrix associated with the graph.

Quantifying the performance of these segmentation techniques can be done with benchmark datasets where human observers have marked the boundaries of the different objects. These reveal that current state-of-the-art algorithms, though much better than previous approaches, still fall short of human level performance. In many applications, such as, for example, in biomedical image analysis, an interactive segmentation approach is used where a human user provides hints to the computer program to complete the segmentation process. The long-term goal is to have fully automated approaches. Such algorithms would likely combine bottom-up processes driven by pixel similarity of brightness, color, and texture with top-down processes based on recognition of familiar objects to arrive at final segmentations of the image or video stream.

Recognition

Humans use vision to recognize people (“Barack Obama”), places (“Times Square”), objects (“chairs”), scenes (“beach”), and actions (“jumping”). Recognition can be of a category (“dog”), or of a specific individual (“my dog Rover”). Many actual and potential applications of computer vision are based on recognition. Examples include identification of humans by face recognition, detection of suspicious activities by surveillance cameras, searching for images of tigers on the Internet, organizing people’s personal photo collections, optical character recognition of ancient manuscripts, or finding a building entrance for a robot.

Recognition is challenging. There is enormous variability in the visual appearance of an object because of factors such as different poses with respect to a camera, differing shading patterns caused by variations in illumination, or for categorization tasks, the variation from one individual to another. Even identification of exactly the same person has to deal with, for instance, differences in clothing, effects of age, or the presence or absence of sunglasses.

Most computer vision approaches to recognition treat it as a problem of machine learning, specifically “supervised learning” or “pattern classification.” Consider the problem of detecting faces in an image, a capability that can assist automatic focusing in digital cameras. To learn to detect blocks of pixels in an image that contains faces, one must train a program to distinguish between positive examples (various images of faces) and negative examples (images of other objects such as chairs, trees, sky, etc.). Such programs are called classifiers, and they take as input a set of “features” derived from the portion of the image corresponding to the object. In the case of face detection, useful features might be the presence or absence of horizontally oriented dark bars roughly around the putative location of the eyes and mouth. The learning program has access to a large number of precomputed features; what it figures out is the weight, positive or negative, to assign to different features. There are a number of different methods for training classifiers such as neural networks, support vector machines, decision trees and nearest neighbors. Choices of features vary, but the most popular, such as the so-called scale invariant feature transform (SIFT) feature proposed by David Lowe, are based on spatial configurations of oriented edges, not too different from the representations that might be produced in the primary visual cortex of primates.

Continuing with the previous example because an image doesn't consist of just a single face, the face may be present at any location at an image, and it could be small or large (corresponding to the imaged person being far or near), the classifier has to examine candidate windows at different possible locations and scales. This is where image segmentation, if successful, could be of great help because it could directly pinpoint a small set of candidate locations and scales. Conversely, recognition, if successful, enables one to mark a region as corresponding to a single object. As mentioned earlier, the right approach will likely be a combination of top-down and bottom-up processing, a subject of much active research in computer vision.

Computer vision techniques for object recognition have a long way to go before they can approach the level of performance of humans. By 2009, leading approaches are able to distinguish among a hundred or more visual categories if objects were present in isolation, but performance degraded considerably when the objects are in scenes with multiple objects potentially occluding each other. Contrast this with humans who, it has been estimated, can distinguish on the order of 30,000 visual categories. There are some clear success stories of computer visual recognition, however: reading printed text, handwritten digit recognition for postal envelopes and bank checks, and face detection in general images. Progress in recent years has been rapid, aided by advances in computing power and the availability of large collections of images for training classifiers.

Reconstruction

The process of image formation, whether in the eye or in a camera, results in the loss of depth information. All points in the external three-dimensional (3-D) world that lie on a ray passing through the optical center are projected to the same point in the two-dimensional image. During reconstruction, we seek to recover the 3-D information that is lost during projection.

Humans use many different cues in the image or video stream, such as binocular disparity, optical flow, position on the ground plane, linear perspective, texture gradients, shading, and occluding contours to achieve this goal. In computer vision, researchers have developed algorithms for exploiting these cues with varying degrees of success.

Of these techniques, the most effective and widely used ones for 3-D reconstruction in computer vision are based on multiple views. For example, we might go around a building and take many photographs of it from different viewpoints. Key points are extracted from the different images, and their correspondences determined; that is, we match key points in the different views that correspond to the same point in the world. When the projections of a sufficient number of points in the world have been matched in multiple images, it is theoretically possible to deduce the 3-D locations of the points as well as the relative position and orientations of the cameras. These techniques have been applied in many different settings. Examples include the compositing of synthetic objects in natural scenes or vice versa for producing special effects in movies, the creation of computer models of real world objects and buildings for populating virtual reality environments, or the accurate 3-D tracking of cricket or tennis balls using multiple cameras for sports broadcasts.

Just as humans can recover 3-D information about a scene from a single image, based on so-called "pictorial" cues, there has also been work in computer vision on trying to exploit cues such as texture, shading, and contour in a single image.

- If the image contains visual texture, then assuming that the arrangement is periodic, or at least regular in a statistical sense, it is possible to recover surface orientation and shape from a single image.
- Shading, or spatial variation in the image brightness, is determined by the spatial layout of the scene surfaces, their reflectance properties, and the arrangement of light sources. If one makes some simplifying assumptions, then a number of techniques can invert the process and recover the surface orientation and shape.

- Even if the image is abstracted down to a line drawing, humans can infer 3-D shape. This results from a combination of cues. For example, upon recognition that the line drawing corresponds to a rectangular block shaped building, one can infer something about the 3-D orientations of the principal planes of the building. If objects are recognized as resting on a ground plane, then their two-dimensional (2-D) position on the image plane conveys their 3-D distance from the viewer. Higher objects, closer to the horizon on the image plane, are further away.

Current thinking on the problem of inference of 3-D structure from texture, shading, and contour tends to be in the framework of Bayesian inference. Given a single image, many possible 3-D worlds could project to the image. It therefore makes sense to choose the one that is the most probable, an insight dating back to Hermann von Helmholtz's notion of perception as "unconscious inference." The world we live in has certain statistical regularities—it contains distinct objects, they often rest on a ground plane, they often move rigidly, and so on. By modeling these in a probabilistic framework, we can start to exploit Bayesian inference methodology. This paradigm has proved to be quite productive in theorizing about human perception, but it should be noted that by 2009 computer vision algorithms for exploiting pictorial cues were still not as general and robust as those based on multiple view analysis.

Controlling Action

Vision is a superb sensor for motor control, and its evolution in animals had much to do with the survival advantages that resulted from it. J. J. Gibson's observation, "We see in order to move, and we move in order to see" captures this succinctly. Autonomously driven robots and cars need sensors for controlling locomotion, and vision and vision-like sensors (e.g., sonar, laser range finders) are obvious choices.

Note that a complete 3-D reconstruction of the external environment, though helpful, is not absolutely necessary. To see this, consider the task of driving a car on a freeway. The driver needs to

- Keep moving at a reasonable speed.
- Control the lateral position of the vehicle in its lane—make sure it stays in the middle and is oriented properly.
- Control the longitudinal position of the vehicle—keep a safe distance from the vehicle in front of it.

The information needed for lateral and longitudinal control is just a small part of that potentially obtainable from a complete 3-D reconstruction of the freeway environment.

This simplifies the task considerably. One can design feedback control laws using visual measurements, and, as far back as the 1980s, Ernest Dickmanns' group in Munich, Germany had demonstrated automated freeway driving using these principles. In 2004, 2005, and 2007, the Defense Advanced Research Projects Agency of the United States organized "grand challenge" races for robot vehicles, the first two in rugged outdoor terrain and the most recent in an urban street environment. No vehicles completed the 2004 race, but several were able to successfully complete the 2005 and 2007 races.

Self-driving cars are not the only examples of the use of visual sensing for motor control. Mobile robots have many present and future applications ranging from autonomous exploration of other planets, operating in hazardous environments, and, of course, the still-awaited household robot that would wash dishes, launder clothes, and vacuum the floor. Unstructured environments such as the home pose the greatest challenges, but potentially the biggest payoffs in settings such as elder care. Although laboratory demonstrations of hand-eye coordination tasks such as reaching, grasping, or insertion, have existed for quite some time, the home robot is still in the distant future.

Past and Future Progress

Computer vision is a field that, like many other areas of artificial intelligence, began in the 1960s in a phase of wildly irrational exuberance and optimism. By the 1970s, researchers had to confront the inevitable disappointment and to realize that the challenges were real and immune to quick fixes. Since then, computer vision has gradually assumed the character of other science-based engineering disciplines. Progress has resulted from geometric and statistical modeling, implementation of ideas and algorithms, accompanied by careful experimentation to determine what works and what does not. The steady advance of computing power, both in processor speed and storage, has helped, as has a healthy exchange of concepts and findings with psychologists and neuroscientists studying biological vision. This has resulted in a large number of successful applications, and progress in the field shows no signs of slowing down.

Jitendra Malik

See also Action and Vision; Bayesian Approach; Depth Perception in Pictures/Film; Object Perception; Perceptual Organization: Vision; Spatial Layout Perception, Psychophysical

Further Readings

Dickmanns, E. D., & Mysliwetz., B. D. (1992). Recursive 3-D road and relative egostate recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 14,* 199–213.

Forsyth, D. A., & Ponce, J. (2003). *Computer vision: A modern Approach.* Upper Saddle River, NJ: Prentice Hall.

Hartley, R., & Zisserman, A. (2004). *Multiple view geometry in computer vision.* Cambridge, UK: Cambridge University Press.

Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision, 60*(2), 91–110.

Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(8), 888–905.

Trucco, E., & Verri, A. (1998). *Introductory techniques for 3-D computer vision.* Upper Saddle River, NJ: Prentice Hall.

CONSCIOUSNESS

Twenty seconds ago, how did it feel in your right big toe? To answer this question, you probably direct your attention to that toe and become conscious of how it feels now. Then you guess that, 20 seconds ago, it probably felt the same. But, did you really have a conscious experience of that toe 20 seconds ago? Or did you have an unconscious experience that only became conscious after you attended? Or was there no toe experience conscious or unconscious before you attended? We are all so familiar with conscious experiences that we take them for granted. But even a simple question about the conscious feeling of a toe can raise perplexing issues. Indeed, consciousness is among the most controversial topics in science and philosophy, as discussed in this entry.

What Is Consciousness?

There is no widely accepted definition of *consciousness.* It is widely agreed, however, that consciousness is not synonymous with *mind,* because there appear to be many unconscious mental states and processes. Consider, for instance, the visual processes that underlie stereovision, which is our ability to see depth using both eyes. The processes involve sophisticated computations of small differences (called disparities) between the images at the left and right eyes, followed by computations of depth and shape from these disparities; we are conscious of the depth and shape, but not of the processes that compute disparities, depth, and shape. The auditory processes that underlie our ability to localize the position of a sound source require sophisticated computations involving differences in phase and intensity of the acoustic signals at the two ears; we are conscious of the location of the sound source, but not of the processes that compute this location. Such examples as these suggest that *most* mental states and processes are unconscious.

A mental state or process is conscious if there is something it is like to be in that state or engage in that process. For instance, what it is like to see a flat plate differs from what it is like to see a round cup, and both are conscious states. What it is like to hear a plane overhead differs from what it is like to hear a mosquito buzz your ear, and again both are conscious states. What it is like to consider a logical argument differs from what it is like to hesitate between pizza or sushi, and both are conscious processes. There is something it is like to see a ruby red grapefruit, endure a pounding headache, enjoy a promising idea, imagine the canals of Venice, feel elated or depressed, or dream that you are flying. Each is conscious or, as it is sometimes put, each has a *phenomenal character* with *phenomenal qualities,* that is, *qualia.*

It is helpful in the study of consciousness to classify conscious states at three levels of detail: boot, general, and specific. The *boot* level simply describes whether a person is conscious at all. Examples of a boot-level transition from the unconscious to the conscious include being awakened from a dreamless sleep, revived with smelling salts after a concussion, or brought up from general anesthesia. The *general* level describes a global quality of consciousness, such as being drowsy, alert, sober, drunk, in a dream, or under hypnosis. The *specific* level describes particular conscious experiences (particular qualia), such as the chirp of a sparrow, the smell of frying bacon, the feel of velvet, or the glisten of dew on a blade of grass.

Approaches to the Study of Consciousness

There are two broad approaches to the study of consciousness: philosophical and scientific. The two approaches interact profitably, but their methods are substantially different.

Philosophical approaches focus on conceptual issues raised by consciousness. These issues are primarily, though not exclusively, ontological and epistemological. The central *ontological* issue is the nature of consciousness and its place in the world. What is consciousness? How is it related to space, time, and matter? The central *epistemological* issue is how we can know whether something is conscious. I think I am conscious, and I believe you are as well. But how can I know for sure that you are conscious? And how do I know that I am conscious? How do I know that a rock is not conscious? These central issues spawn a variety of related issues and philosophical theories.

Scientific approaches construct theories that make specific empirical predictions, and test these predictions in controlled experiments. The theories address several key questions: Why, and how, are some mental states conscious or unconscious? How are the three levels of consciousness related to activity in the brain? These questions are studied from the perspectives of several fields, including cognitive neuroscience, computer science, and physics.

Scientific Approaches

What is the biological basis of consciousness? In 2005, the journal *Science* placed this question second in a list of the top 125 questions that are as yet unsolved. Most scientists who study consciousness are trying to answer this question. The reason for their confidence that there is a biological basis for consciousness is the large number of correlations between neural activity and conscious experience.

There is, for instance, in the occipital lobes of the brain a tight correlation between activity in the primary visual cortex (called area V1) and conscious visual experiences. If area V1 of the left hemisphere is destroyed, then conscious visual experience in the right half of the visual world is also destroyed; the person is blind in this half of the visual world. If smaller portions of V1 are destroyed, then conscious visual experiences in correspondingly smaller portions of the visual world are destroyed. Electrical stimulation of V1 is correlated with anomalous visual experiences.

Activity in visual cortical area V5 is correlated with the conscious experience of visual motion. If area V5 in the right hemisphere is damaged, then the conscious experience of motion in the left visual field is destroyed, a condition called hemi-akinetopsia. Activity in area V5 can be inhibited by appropriate magnetic fields generated by transcranial magnetic stimulation (TMS); this leads to temporary hemi-akinetopsia.

Most of the empirical work on consciousness involves careful study of such correlations, seeking to find, for a specific kind of conscious experience, a specific neural correlate of consciousness (NCC). An NCC for a specific kind of conscious experience is a minimal system within the brain whose activity is normally sufficient for the presence of that conscious experience.

Much of the theoretical work on consciousness involves proposing specific candidates for NCCs. Some proposals point to microscopic properties of individual neurons, such as coherent and incoherent quantum states in neuronal microtubules. Others point to certain brain areas, such as the claustrum, a thin layer of gray matter that lies below and parallel to the cortex. Still others point to patterns of neural activity that loop between the brain's cortex and a subcortical structure called the thalamus. The brain is a complex organ with many levels of organization, from the submicroscopic, to the synaptic, to the neuronal, to local neural networks, to more global interacting neural systems.

Theoretical work on NCCs tries to specify where in the brain a NCC is located and at what level of brain organization it operates. There is not yet consensus about location and level for any NCC.

Proposing candidates for the NCCs is only half of the theoretical enterprise for scientists studying consciousness. The other half is explaining why there are such correlations. If, for instance, an NCC is found for the conscious experience of the taste of dark chocolate, then the next step is to explain why the brain activity in the NCC is correlated with that taste. Does the brain activity *cause* the taste? If so, precisely how? Why does this brain activity cause the taste of chocolate but not, say, the smell of lemon? Why does this brain activity cause any conscious experience at all? There are not yet any scientific theories that try to answer such questions (this is sometimes called, for good reason, the *hard problem* of consciousness). The focus, for now, is on constructing descriptive theories of *what* the NCCs are, rather than explanatory theories of *how* the NCCs actually work.

Brain-Imaging Experiments

The empirical search for NCCs often employs brain-imaging technologies, such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI). EEG measurements of brain activity give excellent temporal resolution, on the order of 1 millisecond, but poor spatial resolution because each electrode reports the summed electrical activity of millions of neurons. fMRI measures brain activity indirectly, through changes in blood flow that are correlated with changes in brain activity; it provides better spatial resolution than EEG does, but poorer temporal resolution.

These techniques can be used, for instance, to search for NCCs associated with shifts in conscious experience during "binocular rivalry." Suppose one presents a homogeneous green field to the left eye and a homogeneous red field to the right eye. Subjects will report that they consciously experience the whole visual world as green for a few seconds, then they report that red starts to invade and eventually spreads until they consciously experience the whole visual world as red. The green and red experiences continue to alternate in an apparent competition for perceptual dominance, hence the name *binocular rivalry*. This is an ideal situation for studying the NCCs associated with a conscious shift in perceived color because the stimuli to the two eyes remain constant even at the moment that the conscious perception shifts. Thus, any brain activity correlated with the shift in conscious perception cannot be dismissed as resulting from a change in the visual stimuli.

Using fMRI, one can determine which brain areas have changes in activation that are temporally correlated with the subject's report of a perceptual shift. EEG can do the same job, but with an interesting twist. One can constantly flash the green stimulus at a certain temporal frequency, say seven flashes per second (7 hertz [Hz]), and at the same time flash the red stimulus at a different frequency, say 14 Hz. The neural systems that process the green stimulus will have activity at many different temporal frequencies, but they will have considerable activity at 7 Hz; similarly, the neural systems that process the red stimulus will have considerable activity at 14 Hz. Thus, this frequency serves as a tag, to let the experimenter know what color a particular neural subsystem is processing. When the subject reports a shift in perceived color, one can then look for brain areas in which there is, at the time of report, a shift in the frequency tag. These areas are candidates for the NCC of the shifts in color experience.

Philosophical Approaches

Is there a relationship between biology and consciousness? This is a key question asked by many philosophers who study consciousness. Notice that this question differs from the question asked by the journal *Science,* namely, What is the biological basis of consciousness? The *Science* question assumes that there is a biological basis of consciousness. The philosopher questions this assumption, and asks whether there is a biological basis of consciousness. Maybe consciousness arises not from biology, but from physics. Or perhaps consciousness and biology exist side by side, neither arising from the other. Or perhaps physics and biology arise from consciousness. There are many possibilities, and philosophers debate their relative merits using logical arguments, thought experiments, and appeals to the best current evidence from science.

One thought experiment that is widely discussed concerns the possibility of philosophical zombies: Is it possible that there could be a person who is atom-for-atom identical to you, who walks, talks and, in every way, behaves just as you would, but who is utterly without conscious experiences? If this zombie were pinched, it would wince and cry just as you would, yet it would experience no feelings of pain.

Some philosophers claim that such zombies are logically possible. This would mean that consciousness is not determined by the physical facts or functional properties of an organism because these facts and properties are shared by the original and its zombie. Instead, consciousness would be something extra beyond the physical and functional. It would also mean that consciousness could not be shaped by natural selection because natural selection can only select among options that make a functional difference to the organism's reproductive fitness; if zombies are possible, then consciousness makes no functional difference, and hence no functional difference in reproductive fitness.

Other philosophers claim that zombies are not logically possible. Once you have specified the physical facts and functional properties of an organism, then all the facts about consciousness are also fixed. Many philosophers in this camp are reductive functionalists, who claim that consciousness is *identical* to certain functional properties. In organisms, these functional properties are most likely properties of the nervous system, but other systems, perhaps sophisticated computers, having internal functional states, could be conscious. According to the reductive functionalist, these functional states do not *cause* consciousness, they *are* consciousness; in the same way, 12 does not cause a dozen, it is a dozen. The reductive functionalist is thus exempt from the duty of explaining how functional properties cause consciousness. Instead, this functionalist must give a principled account of why certain functional states are conscious and others are not. No such accounts are widely accepted yet.

Still other philosophers claim that we are all zombies, and that none of us has any conscious experiences. Our belief that we are conscious is an illusion, perhaps caused by certain computational processes in the brain that monitor the states of other brain processes. Few philosophers are willing to defend this view. Most admit that we are not infallible about our own conscious experiences, and that we can at times be wrong about what we think we are experiencing. But few are willing to make the stronger claim that we are so wrong that we are in fact zombies: If we are that wrong about conscious experiences, is there anything we can be reasonably sure about?

Historical Perspective

The modern philosophical study of consciousness is usually traced to the work of René Descartes (1596–1650), who proposed *substance dualism,* the doctrine that there are two kinds of substance in the universe: physical substance and mental substance. According to Descartes, each human being is a composite of both substances, and they interact in the brain at the pineal gland. Few professional philosophers or scientists now advocate substance dualism, although a version of it was defended by Karl Popper (1902–1994) and John Eccles (1903–1997) in their 1984 book, *The self and its brain.*

George Berkeley (1685–1753) proposed a version of *idealism,* the doctrine that only consciousness exists, and that what we call the physical world is really the contents of conscious experience. Idealism was developed further by Immanuel Kant (1724–1804) and Georg Hegel (1770–1831), and enjoyed substantial popularity in the late 19th century.

Critiques of idealism by Bertrand Russell (1872–1970) and G. E. Moore (1873–1958) contributed to a decline in its influence. In its place, the dominant view became physicalism, the doctrine that all that exists is the physical world and that consciousness is caused by, emerges from, or is a property of certain physical systems. This switch was particularly striking in the field of psychology. The classic 1890 book, *Principles of Psychology,* by William James (1842–1910) freely discusses consciousness and the hard problem of its relationship to brain activity. But the behaviorist psychology of William Thorndike (1874–1949) and B. F. Skinner (1904–1990) restricted itself to the physical vocabulary of stimulus and response and considered consciousness to be an unsuitable topic for scientific study. Behaviorism was superceded about 1960 by more cognitive approaches to psychology, and consciousness once again became a topic for active scientific study about 1990. Most of the current scientific and philosophical work on consciousness assumes that some version of physicalism is true, and takes seriously the hard problem of explaining how conscious experiences could arise in a purely physical world.

Future Research

Empirical investigation of the neural correlates of consciousness is currently the most active area of consciousness research. It promises to remain so, largely because of the complexity of the brain, with its billions of neurons and trillions of synapses, and because of the complexity and variety of conscious experiences.

This will continue to be complemented by theoretical work that tries to explain how neural activity can cause, or give rise to, conscious experiences. Progress on this problem has been remarkably slow, with no theory yet proposed that has the minimal explanatory power expected of a genuine scientific theory. Current theories are sketches, serving primarily to suggest directions where a genuine theory might be found: perhaps quantum states of microtubules, or perhaps looping patterns of neural activity between the thalamus and cortex, or perhaps the informational complexity of neural activity. This promises to be an active area of research, with big new ideas needed for its solution.

The precise relationship between consciousness and attention is controversial and promises to be an active area of research. Are they in fact identical? If not, can there be attention without conscious experience, or conscious experience without attention? Is attending to your right big toe identical to having a conscious experience of that toe, or do you have a conscious experience of that toe even when you are not attending?

Also controversial is the relationship between the self and conscious experience. Does consciousness require a self? Can there be conscious experiences without an experiencer? Can there be an experience of the smell of garlic without a self that experiences that smell? If a self is required for consciousness, why is this the case and what is the nature of the self?

As these examples indicate, research into consciousness is highly multidisciplinary, spanning philosophy, neuroscience, cognitive science, computer science, and physics. The insights that emerge will transform our understanding of human consciousness, and therefore of human nature and its place in the world.

Donald D. Hoffman

See also Attention and Consciousness; Mind and Body; Private Nature of Perceptual Experience; Unconscious Processes

Further Readings

Blackmore, S. (2004). *Consciousness: An introduction.* New York: Oxford University Press.

Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory.* New York: Oxford University Press.

Crick, F. (1995). *The astonishing hypothesis: The scientific search for the soul.* New York: Scribner's.

Edelman, G., & Tononi, G. (2000). *A universe of consciousness.* New York: Basic Books.

Hoffman, D. (2008). Conscious realism and the mind-body problem. *Mind & Matter, 6,* 127–181.

Koch, C. (2004). *The quest for consciousness.* Englewood, CO: Roberts & Company.

CONSCIOUSNESS: DISORDERS

As far as we know, a living being without a brain cannot have conscious experiences. But are certain parts of the brain more critical for conscious awareness than others? More specifically, how might disorders of perceptual consciousness after damage to the human brain inform the scientific world about relationships between neural systems and conscious perceptual awareness? In many cases, the loss of awareness is accompanied by indications that "invisible" information is nevertheless still being processed by the visual cortex. Studies of neuropsychological disorders of perceptual consciousness have shown that the primary visual cortex is neither necessary nor sufficient to account for perceptual awareness. This entry describes blindsight, unilateral visual neglect, Bálint's syndrome, visual agnosia, and further issues relating to consciousness disorders.

Blindsight

Blindsight refers to the unconscious detection of features that are not consciously perceived because of damage to the primary visual cortex. It was first reported in 1977 by Larry Weiskrantz. He collected evidence from a man who had undergone neurosurgery to remove a tumor from his primary visual cortex. The surgical procedure left the patient blind in the visual field opposite the affected hemisphere (contralateral). The primary visual cortex is the cortical entry point for most of the visual information received from the eye. It is located at the back of the brain with the right primary cortex receiving visual information from the left visual field (to the left side of fixation) and the left primary cortex receiving visual information from the right visual field (to the right side of fixation). The patient Weiskrantz studied lost his right primary visual cortex and saw nothing to the left of where he was currently looking, although, if the visual scene remained stable, he was able to compensate for his visual loss by moving his eyes. Thus, to probe his remaining visual abilities, he was tested by presenting stimuli briefly before he had time to move his eyes, thereby moving the stimulus into his good visual field.

Weiskrantz showed that after a light was briefly flashed in the patient's bad visual field, he was able to move his eyes to its previous location, or point at it much more accurately than if he had been guessing. Yet guessing is exactly what he reported he was doing; consciously, he saw nothing. Further testing revealed that he could discriminate the orientation of simple lines as well as the direction of moving dots above chance levels, although again, he denied seeing anything. Weiskrantz called this condition blindsight. At first, questions were raised, attributing the results, for example, to light scatter into good areas of the visual field, remaining pockets of the cortex, or undetected eye movements. However, in most patients, these alternative accounts have been ruled out.

Accepting that blindsight is a genuine phenomenon, what implications follow for the neural basis of consciousness? Does blindsight mean that the primary visual cortex is necessary for perceptual awareness of visual information? Other pathways for visual information to the cortex are known (e.g., superior colliculus to the frontal eye field; pulvinar and its connections to the posterior visual association cortex), but the fact that blindsight results from lesions of the primary visual cortex suggests that these other pathways are (at least initially) insufficient to support conscious visual awareness. One interesting question is whether other visual pathways can be trained to mediate conscious experience. Some patients with primary cortex lesions have shown spontaneous recovery of conscious vision, even after years of blindness, and monkeys, after full ablation of the primary visual cortex, have been trained to show improvements in visual discrimination. In a recent review, Tony Ro and Robert Rafal suggest that this evidence, at least in part, reflects learned access to visual awareness through pathways from the superior colliculus to cortical areas. Animals with lesions of both the visual cortex and the superior colliculus do not regain vision after training. These findings imply that visual awareness is at least partly a learnable skill. Normally, the salient information coming from the primary cortex dominates conscious visual experience, but, when that is removed, information normally used for other purposes (e.g., the planning of eye movements) can eventually create conscious experiences. One question that remains unanswered is the degree to which regained conscious experience is similar to normal conscious experience.

Blindsight has also been studied to address more cognitive questions, such as whether selective attention can influence unconscious processes. Some argue that attention and consciousness are so closely related that they may be the same thing. For instance, ignoring an item in a visual scene is an active process, but there is continued awareness that that item still exists and can become the focus of attention on demand. But if selective attention could also affect unconscious perception in blindsight, consciousness and attention would be separable processes. In a study by Robert Kentridge, a patient's attention was drawn to a particular location in his blind field (by an arrow in his seeing field pointing to a blind field location). The patient was faster to correctly guess the orientation of an unseen line when the cue pointed to the location of the line than when it pointed elsewhere. Thus, attention facilitated his response to an unconscious stimulus, ruling out the hypothesis equating attention with consciousness

Unilateral Visual Neglect

Blindsight patients see nothing in the area corresponding to the damaged area of the cortex, but they still understand that that area of space exists. A quite different type of problem, known as unilateral neglect, occurs when damage includes the parietal lobe (usually in the right hemisphere). In these cases, it is as if the left side of the world no longer exists, either because it is not adequately represented in an internal spatial framework or because it is no longer attended. Neglect patients ignore objects in the left visual field, even a $20 bill waved in front of them! They may leave the food on the left side of their plate or shave only the left side of their faces. They draw only half of an object (e.g. a clock, a tree, or a house). Neglect can affect other sense modalities besides vision. One woman hopped around on one leg even though her other leg worked fine. When asked why, she said, "How else can I get around?" Although blindsight shows that the primary visual cortex is necessary for conscious experience (at least in the early stages of blindness), unilateral neglect shows that it is not sufficient to support conscious perceptual experience.

As in blindsight, several studies have shown that undetected visual information presented in the neglected field can influence behavior. For instance, studies have shown that stimuli presented in the "bad" field on one trial (e.g., a drawing of a baseball bat) will prime (speed up) responses to a related stimulus presented subsequently in the good field (e.g., the word *sport* compared with *banana*). Jon Driver and his colleagues have shown that the organization of stimuli placed in the bad field can affect how a pattern is perceived in the good field. For example, "unseen" cues to depth can influence the perceived depth on the good side. Failure to detect stimuli in these cases means that the patient reports seeing nothing there at all.

A straightforward account of unilateral neglect is that attention can no longer move to the side of space opposite the damage. But the straightforward explanation runs into difficulty. For instance, when asked to copy the drawing of a plant with two flowers (one on the left and one on the right) planted in a single pot, a patient with left neglect drew the right half of the picture, leaving out half the pot and the left branch completely, as he would if neglecting the left side of space. But when the experimenter removed the pot and the shared stem by cutting off the bottom part of the picture, the patient drew the right side of each flower. In other words, he neglected the left side of each separate object, not the left side of space. If you move a picture behind a slit so that attention and vision are focused on a single narrow area of space, patients still neglect the left side of the object, although it is distinguished from the right only by being visible at a different time. So the brain converts time into space and then neglects the left side of the reconstructed image! To do this the brain must first "see" the whole object so that the brain can then suppress its left side.

There are some other exceptions to the claim that all stimuli presented on the neglected side are undetected. With other colleagues, Lynn Robertson showed that a unique feature, like a red circle in an array of many green and blue items, can be relatively easy to detect on the "neglected" side, whereas combinations of two features, such as a red circle among green circles and red triangles (i.e., a unique conjunction of features) is not. These findings are consistent with a theory of attention that Anne Treisman proposed in the late 1970s (feature integration theory). She suggested that unique features could be detected without attention or awareness of the spatial locations of those features, but conjunctions of two or more features required focused attention to their locations. For example, when looking for one's blue car in a parking lot, detecting that there is something blue in the visual scene does not require knowing where the blue is located, but identifying it as a blue car does. The data from patients with unilateral neglect demonstrate that the intact primary visual cortex is sufficient to support the perception of unique features, but that focused attention, controlled by the damaged hemisphere, is needed when these features must be bound together.

Is there a conflict here with the priming effects described earlier? Wouldn't the (unconscious) perception of a baseball bat require the integration of lines, curves, and angles? Perhaps the integration of different dimensions such as color and shape requires spatial awareness but the integration of curves, lines, and angles to form an object does not? If so, are there spatial maps below the level of awareness that support some types of integration but not others, and why are these not sufficient for conscious access of their content? Results from patients with spatial awareness deficits suggest that such maps exist, but research is needed to understand their nature and what neural mechanisms can be engaged to access them for consciousness.

Bálint's Syndrome

Unlike unilateral neglect, which results from damage to one side of the brain, Bálint's syndrome is caused by parietal damage to both sides. This syndrome is rare, and the experimental literature is generally limited to single case studies. Nevertheless, findings are generally consistent. Patients with Bálint's syndrome are functionally blind in both visual fields. When the syndrome is severe, patients left on their own will bump into walls and objects in the room, or ignore a fork at the side of a plate. They are unable to navigate from one room to another within their own homes or to do anything that requires accurate spatial information in the external world. When tested on spatial discriminations, they are unable to locate an object, whether through reaching, pointing, or verbally reporting. This is not a loss of the meaning of spatial terms because the patients can still locate touches on their own bodies.

Patients with Bálint's syndrome are not completely blind. They can see a single object, which may or may not be the object at which they are looking, but they act as if other objects are not present. Another puzzling fact is that the one object they see is usually seen as a whole, regardless of whether it is complex or simple, large or small. The famous Russian neuropsychologist Alexander Luria drew two circles side by side on a sheet of paper and asked a patient with Bálint's syndrome to tell him what he saw. The patient reported seeing one circle but could not see the other, even when prompted. Luria then showed the same two circles connected by a line and asked the same question. The patient now reported seeing "spectacles." Connecting the two circles changed his perception from one circle to one pair of eyeglasses.

Although it was originally thought that the same area of the parietal lobe was critical in producing unilateral neglect and Bálint's syndrome (the only difference being a lesion in one versus two hemispheres), reviews of the literature suggest that different areas of the parietal lobe are involved. Unilateral visual neglect is most often associated with parietal areas (supermarginal gyrus) that are somewhat in front of those associated with Bálint's syndrome (angular gyrus). This anatomical distinction is accompanied by differences in visual experience. Although patients with unilateral neglect miss parts of objects on the neglected side, patients with Bálint's syndrome perceive a whole object but lose its spatial relationship to other objects. Also, patients with unilateral neglect usually report the location of objects accurately once they detect them, but those with Bálint's syndrome cannot. Both syndromes disrupt spatial awareness, but in different ways.

A disruption of spatial awareness means that spatial attention is also disrupted in both syndromes. It follows that feature integration should be deficient, and single-feature pop-out should remain relatively intact, and research has confirmed this prediction. The Bálint's patient that Robertson and Treisman studied also made many "binding" errors. For example, if they showed him a red X and a blue T, even for as long as 10 seconds, he frequently reported seeing either a red T or a blue X. However, again evidence indicates that the deficits may be primarily in conscious experience. Studies using priming measures demonstrated that the locations of unique features were processed preconsciously, although they remained inaccessible to perceptual experience. M. S. Kim and Robertson instructed a Bálint's patient to make a speeded response whenever an arrow appeared at the center of the screen where he was looking. Unbeknownst to the patient, a feature display appeared simultaneously in the periphery. When the arrow pointed to the unique feature, response times were faster than when it pointed to one of the other items in the display.

Visual Agnosia

Whereas the visual pathways through the dorsal areas of the human cortex, including the parietal lobe, appear to deal primarily with spatial representation, the pathways through ventral areas, including the temporal lobes, seem more involved with object perception. Lesions in the occipital and temporal lobes can produce striking deficits in awareness and recognition of objects. These deficits are known as agnosia. Patients with agnosia are not blind—they see the visual features of the objects present but they are unable to make sense of them or to recognize what they see. One such patient with lesions in both temporal lobes was studied by David Milner and Melvin Goodale. When they showed her a paper card and asked her what it was, she was unable to say that it was a card or to see which way it was oriented. Yet when they presented a slit, like in a mailbox, and asked her to post the card, she immediately rotated it to the correct angle to pass through the slit. Her hands seemed to "know" something that her mind did not. The neurologist Rafal filmed another patient with visual agnosia who was unable to recognize objects. When the patient was shown a picture of a clarinet, he hesitated in naming it, suggested it was a "pencil," but meanwhile his fingers began to play an imaginary clarinet. These are cases where a patient seems both to see something and not to see it—to know something and not to know it at the same time.

Are these fragmentations of knowledge just some peculiarity of damaged brains, or do normal people show any comparable separation between knowledge for conscious perception and knowledge for action? Goodale showed that some visual illusions that affect conscious perception of objects might have no concomitant effect on motor interactions. For example, an illusion that induces conscious perception of increased size has no effect on the size of the gap between observers' fingers when they pick up the object. As in the patients, the brain areas that control grasping seem to use different information from those that determine what is consciously seen.

Further Issues

Neuropsychological disorders of spatial awareness are also relevant to philosophical debates. Immanuel Kant argued that there could be no intuition about the external world without the cognitive constructs of space and time. The research discussed in this short essay demonstrates that spatial concepts can disappear, yet simple features can still be perceived. The impact of the loss of space on the control of spatial attention also has implications for understanding object perception. Conscious representations normally bind features into integrated wholes, but without spatial attention, the binding can be lost, and only the features are correctly perceived, supporting a particular role for spatial attention in perceptual awareness.

Damage to higher levels of the ventral visual pathway yields new puzzles about consciousness. We think of ourselves as unitary identities with single organized minds. We either see the clarinet or we don't. Yet, evidence from both patient and healthy populations casts doubt on this unitary notion of individual cognition. The metaphor of an organization seems to fit much better with these dissociations between parts knowing and other parts not knowing particular things. Like the CIA, the brain seems to have information floating around in different offices and different files, not all of it accessible to every subsystem or usable for every purpose. Shapes can be used to shape one's fingers, while being verbally described inaccurately—or the reverse. The information that becomes conscious may be the information that makes a coherent organized story about what is happening, bringing together selected relevant facts from many different specialized subsystems, but leaving the detailed evidence for those conclusions and the detailed implementation of the selected actions to unconscious processors.

Lynn C. Robertson and Anne Treisman

See also Agnosia: Visual; Neuropsychology of Perception; Visual Processing: Primary Visual Cortex

Further Readings

Driver, J., & Vuilleumier, P. (2001). Perceptual awareness and its loss in unilateral neglect and extinction. *Cognition, 79,* 39–88.

Eglin, M., Robertson, L. C., & Knight, R. T. (1989). Visual search performance in the neglect syndrome. *Journal of Cognitive Neuroscience, 4,* 372–381.

Kim, M. S., & Robertson, L. C. (2001). Implicit representations of visual space after bilateral parietal damage. *Journal of Cognitive Neuroscience, 13,* 1080–1087.

Milner, A. D., & Goodale, M. A. (2006). *The visual brain in action* (2nd ed.). New York: Oxford University Press

Ro, T., & Rafal, R. (2006). Visual restoration in cortical blindness: Insights from natural and TMS-induced blindsight. *Neuropsychological Rehabilitation, 16,* 377–396.

Robertson, L. C. (2004). *Space, objects, minds and brains.* New York: Psychology Press.

Robertson, L. C., Treisman, A., Friedman-Hill, S., & Grobowecky, M. (1997). The interaction of spatial and object pathways: Evidence from Bálint's syndrome. *Journal of Cognitive Neuroscience, 9,* 295–317.

Shipp, S. (2003). The functional logic of cortico-pulvinar connections. *Philosophical Transactions of the Royal Society of London, B,* 358, 1605–1624.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Weiskrantz, L. (1986). *Blindsight: A case study and implications.* New York: Oxford University Press.

CONSTANCY

An important function of perceptual systems is to accurately determine *object properties*—physical properties of objects that remain constant. For example, an object's shape, size, and reflectance (fraction of light and spectral distribution of light reflected from the object) typically remain constant and so are considered properties of the object. Examples of object properties (with the object properties *italicized*) are a *brown* cat, a *six-foot-tall* person, a *rectangular* door. *Perceptual constancy* refers to the perception of these object properties as remaining constant even when things such as illumination, distance, and viewing angle cause changes in the stimulation these objects present to the receptors. For example, as a rectangular door swings open, the image it casts on the retina changes from rectangular, when the door is viewed straight on, to a family of trapezoidal shapes, as the door is viewed at different angles.

Perception of the door as remaining rectangular even as viewing angle changes is an example of shape constancy. The ability to perceive properties of objects as unchanging in the face of changing stimulation is an important adaptive property of perception, because it is important for organisms to be aware of the stable properties of the environment. The world would be a confusing and difficult place to cope with if we perceived everything as constantly changing. The person or animal that perceives a charging lion as small when it is in the distance, and as becoming large only when it is upon them, is not likely to survive to perceive another lion! This entry describes perceptual constancy by describing examples of constancy in vision, touch, and olfaction.

Size Constancy

Size constancy occurs when an object's perceived size remains constant even though the object is viewed from different distances. For example, as a person walks away from an observer, the size of the person's image on the retina gets smaller, but the person don't appear to shrink as he walks away, except perhaps at extreme distance. A 6-foot-tall person is judged to be about 6-feet tall when viewed at a distance of 3 feet and when viewed from 12 feet, even though image on the observer's retina is only a quarter as large at 12 feet.

One explanation for size constancy is a "taking-into-account" mechanism that proposes that the perceptual system takes distance into account. The idea of a link between size constancy and distance perception has led to the proposal that size constancy is based on a mechanism called size-distance scaling that operates according to the equation $S = K (R \times D)$, where S is the object's perceived size, K is a constant, R is the size of the retinal image, and D is the perceived distance of the object.

According to the size-distance equation, as a person walks away from an observer, the size of the person's image on the observer's retina, R, gets smaller, but the observer's perception of the person's distance, D, gets larger. Because these two changes balance each other, the net result is that size perception remains constant. This explanation for size constancy is also called the size-distance invariance hypothesis.

A classic experiment by Albert Holway and Edwin Boring investigated the effect of perceived distance on size perception by having observers adjust a comparison stimulus to match the sizes of test stimuli presented at different distances, and by manipulating observers' ability to perceive the distances of the test stimuli. When observers were able to perceive distance accurately, they adjusted the size of the comparison stimulus to correspond to the physical size of the test stimulus, no matter what the distance. In other words, they accurately perceived the sizes of stimuli presented at different distances. However, when information for distance was eliminated, so it became more difficult to accurately perceive the distances of the test stimuli, estimates of the sizes of the test stimuli became less accurate. The conclusion from this experiment and others like it is that size constancy is most likely to occur when distance can be accurately perceived.

Another source of information that can result in accurate size perception independent of the size of the retinal image is relative size—comparing one object to a familiar object with known size. For example, seeing a person standing next to the door to a room will usually (unless the door is atypically small or large) provide information about the person's height. A related example, which also involves comparison to the surroundings, occurs as a person walks away from an observer inside a room. If the person is 6-feet tall and the room has 12-foot high ceilings, the person's height relative to the height of the ceiling remains a constant fraction of the height of the ceiling (6/12 = 0.5 in this example), no matter where the person is in the room. (See Table 1 for a summary of conditions that are involved in size constancy, as well as for the other types of constancy described in this entry.)

Shape Constancy

Shape constancy occurs when perceived shape remains constant as viewing angle changes, as in the example of the door viewed from different angles. Robert Thouless called this constancy of shape perception "phenomenal regression to the real object." The explanation for shape constancy is similar to that for size constancy—observers consider the object's slant. This idea is supported by the results of experiments that show that eliminating information for slant makes it more difficult to accurately judge an object's physical shape.

Lightness Constancy

Lightness constancy occurs when the perceived lightness of a surface (where lightness varies along the dimension, white, gray, black) remains constant as the intensity of illumination is changed. For example, a black Labrador retriever appears black inside a room and still appears black when it is outside in sunshine, even though far more light is reflected from its coat when in sunshine.

When lightness constancy occurs, the perception of lightness is determined not by the illumination hitting an object, but by the object's reflectance. Objects that look black reflect about 5% of the light. Objects that look gray reflect about 10 to 70% of the light (depending on the shade of gray),

Table 1 Perception and Stimulus Conditions for Different Types of Constancy

What remains constant	*What is varied*	*Change in receptor stimulation*
Size	Distance	Image on retina gets smaller as distance increases
Shape	Slant/viewing angle	Image on retina changes shape as object is viewed from different angles
Lightness	Intensity of illumination	Stimulation of receptors increases as illumination increases
Color	Spectral distribution of illumination	Ratio of activation of different cone receptors changes when spectral distribution of light changes
Roughness	Speed of scanning with the fingers	More texture is scanned per unit time as scanning speed increases
Olfactory intensity	Magnitude of sniffing	More molecules of odorant stimulate olfactory receptors as sniff magnitude increases

and objects that look white reflect 80 to 95% of the light. Thus, when lightness constancy occurs, our perception of an object's lightness is related not to the *amount* of light that is reflected from the object, which can change depending on the illumination, but on the *percentage* of light reflected from the object, which remains the same no matter what the illumination.

One explanation for lightness constancy is that when illumination changes, the ratio of the amount of light reflected from neighboring objects in a scene remains constant. For example, if the print on the pages of a book reflects about 5% of the illumination and the white paper of the page reflects 90%, then if the illumination is relatively dim, as occurs inside a room (say 1,000 for the purposes of this example), then the print reflects 50 units of light and the white page 900. However, when the illumination is high, as occurs when in bright outdoor illumination (for example, when illumination is 100,000), then the print reflects 5,000 units of light and the white page reflects 90,000 units. Note that the black print reflects more light in the bright illumination than the white paper reflects in dim illumination, yet the print still looks black. This constancy of lightness perception can be attributed to the fact that the ratio of light reflected from the print and paper remains constant at 5/90 = 0.55, no matter what the illumination.

Another mechanism that plays a role in lightness constancy is the observer's awareness of the conditions of illumination. For example, lightness constancy breaks down when an object is illuminated by a hidden light source. In a situation such as this, a piece of black paper (that reflects 5% of the light) can appear light gray or white if it is the only object visible and is illuminated by a hidden source. When this occurs, the hidden illumination is causing an erroneous perception of a change in object property—normally black paper is perceived as white.

Experiments have also shown that the perception of lightness can be influenced by the interpretation of how light is falling on a three-dimensional object. For example, if light appears to be illuminating one side of a building and the other side is shaded, observers can account for the illumination conditions and so avoid perceiving the more intensely illuminated side of the building as having a different object property (being constructed with lighter colored bricks, for example) than the shaded surface is.

Color Constancy

Color constancy occurs when the perceived color of an object (where color includes perceptions such as blue, green, yellow, orange, red) remains constant even when the spectral distribution of light is changed. For example, a green sweater continues to appear green when illuminated by an old-fashioned light bulb that has a yellowish hue because it is relatively rich in long-wavelengths, and appears green when viewed in daylight, which has an even distribution of wavelengths across the spectrum. The wavelengths actually reflected from the sweater depend on both the reflectance curve of the sweater—green objects typically reflect more light in the middle of the spectrum—and the wavelength distribution of the light illuminating the sweater, but when color constancy occurs perception is determined primarily by the object's reflectance curve.

As for lightness constancy, color constancy has been shown to be affected by the availability of information indicating the nature of the illumination, and color constancy is most robust in environments or displays that contain a number of objects with a range of different spectral reflections, and when the illumination contains a wide range of wavelengths, as normally occurs under normal indoor or outdoor illumination. An example of a breakdown in color constancy is the distorted colors perceived when objects are viewed under lights such as sodium-vapor lights that emit only a narrow band of wavelengths.

Color constancy is enhanced by a process called selective adaptation. For example, when objects are illuminated by a light bulb that emits light with relatively more long wavelengths (so the bulb looks yellowish), the long wavelength illumination adapts or decreases the responses of retinal receptors that are sensitive to long wavelengths. This decreased sensitivity to long-wavelength light acts to prevent objects illuminated by this light from taking on a yellowish hue.

Another mechanism of color constancy is memory color—the effect of memory for the color of familiar objects on color perception. Research has shown that perception of the colors of familiar objects such as a red stop sign or a green tree, are richer and more saturated than the colors of unfamiliar objects that reflect the same wavelengths.

Roughness Constancy

Moving the fingers over a textured surface such as sandpaper results in the perception of roughness. Roughness constancy occurs when the roughness of a surface perceived by scanning the fingertips across a surface remains constant even when the speed of scanning is varied. Experiments in which roughness is estimated for different scanning speeds show that roughness for coarse textures (but not for fine textures) is relatively unaffected by the relative speed of motion between a surface and the hand or fingers. This occurs even though high scanning speeds result in more units of texture being sampled per unit of time, and faster firing rates of neurons that fire to textures. The exact mechanism responsible for roughness constancy is not clear, but two possibilities are that even when the speed of scanning changes, the relative amount of skin deformation may remain the same for a particular roughness. It is also likely that observers are aware of moving the hand at different speeds, and so can account for the speed of movement.

Odor Intensity

Constancy of odor intensity occurs when the magnitude of an olfactory stimulus is judged to remain the same whether the odorant is sniffed gently or vigorously. This was demonstrated by Robert and Martha Teghtsoonian and coworkers, who first showed that when subjects are presented with different concentrations of odorants, they rate the higher concentration as smelling stronger. This result is consistent with the idea that perceived odor is a function of the concentration of odorant molecules that are available to the receptors. However, constancy was then demonstrated by instructing half of the subjects to sniff a particular concentration weakly and half to sniff it strongly. When subjects did this, the strength of sniffing had no effect on their estimates of odor intensity. Thus, for a given odorant concentration, subjects produced the same estimate for weak sniffs and for strong sniffs. This occurred even though more odorant was delivered to the receptors for the strong sniffs compared with the weak sniffs and suggests that subjects are taking sniff intensity into account.

The ability to take sniff intensity into account is adaptive because it increases the chances that subjects will be able to accurately sense the concentration of an odorant, without being misled by different sniff intensities that they might use in different situations. They can also regulate sniff intensity to avoid excessive intake of hazardous odors. By sniffing strongly to detect weak odorants and weakly to sample strong, possibly toxic, odorants (which might pose a hazard if sniffed too vigorously), they can still tell that one odorant is weak and the other is strong.

Degree of Constancy

Constancy is not always complete. For all of the examples described, changes in perception of object properties can occur in response to changes in receptor stimulation. For example, changing the illumination on a piece of paper that looks white in bright illumination might cause it to look slightly grayer. However, the change in perception is generally less than would be expected from the magnitude of the change in receptor stimulation, an effect called *partial constancy*. Even though constancy can be "partial," it still serves an important adaptive purpose, by keeping our perception of object properties relatively constant and preventing our perceptual world of shapes, sizes, lightness, colors, and textures from changing every time the conditions of stimulation change.

E. Bruce Goldstein

See also Color Constancy; Color Perception; Contrast Perception; Object Persistence; Olfaction; Perceptual Development: Visual Object Permanence and Identity; Texture Perception: Tactile

Further Readings

Adelson, E. H. (1999). Light perception and lightness illusions. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 339–351). Cambridge: MIT Press.

Barnes, D. C., Hofacer, R. D., Zaman, A. R., Rennaker, R. L., & Wilson, D. A. (2008). Olfactory perceptual stability and discrimination. *Nature Neuroscience, 11,* 1378–1380.

Delahunt, P. B., & Brainard, D. H. (2004). Does human color constancy incorporate the statistical regularity of natural daylight? *Journal of Vision, 4,* 57–81.

Gilchrist, A. L. (Ed.). (1994). *Lightness, brightness, and transparency.* Hillsdale, NJ: Lawrence Erlbaum.

Gregory, R. L. (1997). *Eye and brain: The psychology of seeing.* Princeton, NJ: Princeton University Press.

Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience, 9,* 1367–1368.

Holway, A. H., & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *American Journal of Psychology, 54,* 21–37.

Lederman, S. J. (1983). Tactual roughness perception: Spatial and temporal determinants. *Canadian Journal of Psychology, 37,* 498–511.

Mainland, J., & Sobel, N. (2006). The sniff is part of the olfactory percept. *Chemical Senses, 31,* 181–196.

Teghtsoonian R., Teghtsoonian, M., Berglund, B., & Berglund, U. (1978). Invariance of odor strength with sniff vigor: An olfactory analogue to size constancy. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 144–152.

CONTENT OF PERCEPTUAL EXPERIENCE

Suppose a subject recognizes George Washington on a dollar bill. It is an open question whether this is part of the subject's perceptual (in this case, visual) experience that the person depicted is George Washington, or whether all that is visually experienced are colors and shapes, with the judgment that it is George Washington occurring farther downstream (later in the perceptual system). According to *Thin* views about the contents of visual experience, these contents are limited to color, shape and illumination. According to *Rich* views, the contents of visual experience can involve more complex high-level features in addition to color, shape, and illumination, such as personal identity (George Washington), kinds (banana, truck), semantic properties (meanings of words), and causation. Rich views thus posit more informational richness in the contents of visual experience than do Thin views. If Rich views are correct, then even if some visual experiences are thin (such as the experience of looking out of an airplane window into a clear blue sky, with nothing else in your peripheral visual field), visual experiences *can* represent that someone is George Washington, or that an object is a truck, or that the ball's collision with the apple caused the apple to move. If Thin views are correct, in contrast, then visual experiences can represent only that there is a layout of colored shapes, variously illuminated, where some part may be moving.

The debate between Thin and Rich views affects the search for neural correlates of visual consciousness. Antecedently held assumptions about whether visual experience is Rich or Thin may influence what researchers are prepared to count as a neural correlate of visual consciousness. If contents are Thin, then the neural correlates are more likely to be limited to early visual areas, such as V1 or V5. If contents are Rich, then the neural correlates are more likely to involve later areas, such as the inferotemporal cortex and the fusiform face area.

The debate between Rich and Thin views interests psychologists because it is relevant to what the neural correlates of visual consciousness are, and on what cognitive processes are inputs to the underlying structures that give rise to visual consciousness. It is also relevant to the structure of disorders such as agnosia, in which subjects can see ordinary objects but cannot recognize them. If the Rich view is true, then agnosia may involve impoverished visual consciousness. If the Thin view is true, then agnosia may be a disorder downstream of visual consciousness, rather than a disorder involving visual consciousness itself. Finally, the debate is relevant to the cognitive structure of delusions such as Capgras syndrome, in which patients take a loved one to be an impostor. If a Rich view is true, so that visual experiences may represent such properties as "being an impostor," then Capgras syndrome may be a normal response to an abnormal experience, consisting in taking things to be the way they appear in the (abnormal) experience. If a Thin view is true, in contrast, whether or not the visual experience of Capgras sufferers is itself abnormal, the response to the experience almost certainly is abnormal. In that case, the disorder would consist of the abnormal response to evidence, perhaps in addition to an abnormal evidential base.

To philosophers, the debate between Rich and Thin views is of interest because of its impact on epistemological issues concerning the evidence provided by perception. If a Rich view is true, then there are likely to be top-down influences on visual consciousness by mental states preceding perception, such as emotion, mood, desire, and beliefs. If such influences are pervasive, then the idea that our perceptual inputs are unaffected by antecedent mental states is probably false. Although some top-down influences may help us perceive what we would otherwise be blind to (as when a radiologist sees a tumor on an X-ray), others may be epistemically pernicious (as it would if a depressive's expectation that everything will look gray influences her experience, so that indeed everything ends up looking gray).

The case of causation illustrates both the debate between Rich and Thin views, and the difficulty in settling it. Whereas a Rich view allows that visual experience can represent that the ball's movement causes the apple's movement, a Thin view will say that only succession of movements is represented in visual experience. In a series of experiments published in the 1960s, Albert Michotte showed adults a range of scenes that elicited causal descriptions. Many of these scenes consisted in "launching": an object A moves toward a stationary object B, makes contact with B, and immediately afterward B starts moving in the same direction as A was moving at the time of contact (see *Causality,* Figure 1). Through an extensive series of experiments involving a wide range of stimuli (including shadows, moving lights, marks on paper, and hefty objects), Michotte tried to isolate the parameters of motion that led adults to classify what they saw as causation. He discovered that even when A or B or both are lights or shadows, adults report that object A *launches* object B exactly when the separate motions of A and B are consistent with a single motion that is transferred from A to B. This led him to posit an "impression of causation" because subjects knew that shadows and lights cannot launch anything.

Though these results are suggestive, they do not conclusively favor Rich views over Thin ones. To get from Michotte's results to a Rich view, what's needed is a principle linking verbal reports to the contents of visual experience. That the verbal report mentions causation does not show that the visual experience represents causation, without some further principle indicating that the reports are reports of experiences, rather than reports of a mental state downstream of experience. For all Michotte's results show, visual experiences may represent the spatiotemporal parameters of motion that elicit causal reports, rather than representing causation itself. If so, then experiences in the launching condition when the stimuli are lights and shadows will be accurate, even though there is no causation. In contrast, if the experiences represent causation (and not just the parameters that elicit causal reports), such experiences would be inaccurate, whether or not subjects would go on to infer that A launched B.

The issue between Rich and Thin views arises for auditory as well as visual experience. We can often recognize who is speaking by their voices. Rich views allow that auditory experience can represent that Frank Sinatra is singing, whereas Thin views allow that it can only represent the sounds and their qualities. The judgment that the person singing is (or sounds like) Frank Sinatra must come downstream of auditory consciousness. Rich views allow that auditory experience can represent that a truck is going by, whereas Thin views allow that it can only represent a gradually fading groaning noise. Rich views allow that auditory experience represents that a string of sounds was a meaningful sentence, whereas according to Thin views it can only represent that a string of sounds.

Like visual experiences, auditory experiences represent the world outside the perceiver, and thus can be accurate or inaccurate. Relevant to this, auditory and visual experiences have spatial content: We hear sounds as coming from different directions, and we see things as having specific spatial locations. Whether a standard visual experience such as the one had when viewing a cube is accurate depends on whether the cube has the properties (including spatial properties) that the visual experience represents it as having. Whether standard auditory experiences such as the ones had when listening to a lecture are accurate depends on whether the lecturer made the noises that the subject auditorily experiences her as making. The exact accuracy conditions of the experiences thus depend on whether Rich or Thin views are true. For instance, suppose you are talking to an impostor on the telephone, but the disguised voice fools you and you wrongly take the impostor to be your spouse. There is clearly an error at the level of judgment, but is there an error at the level of auditory experience as well? If a Thin view is true, then your auditory experience may be correct in this case because it represents the tone of voice that the impostor actually has. In contrast, if a Rich view is true and your experience represents that the person you're hearing is your spouse, then your error will reach all the way down to the auditory experience itself.

Susanna Siegel

See also Consciousness: Disorders; Intentionality and Perception; Philosophical Approaches; Vision: Cognitive Influences

Further Readings

Bayne, T. (forthcoming). Perception and the reach of phenomenal content. *Philosophical Quarterly.*

Byrne, A. (2009). Experience and content. *Philosophical Quarterly.* DOI: 10.1111/j.1467–9213.2009.614.x

Matthen, M. (2009). The feeling of presence. In N. Gangopadhyay, M. Madary, & F. Spicer (Eds.), *Perception, action and consciousness.* Oxford, UK: Oxford University Press.

Siegel, S. (2005). Which properties are represented in perception? In T. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 481–503). Oxford, UK: Oxford University Press.

Siegel, S. (2009). The visual experience of causation. *Philosophical Quarterly.* DOI: 10.1111/j.1467–9213.2008.607.x

CONTEXT EFFECTS IN PERCEPTION

When Groucho Marx observed, "Time flies like an arrow; fruit flies like a banana," he made two essential points about the effect of context on perception. The first point is that an isolated stimulus, such as the word *flies,* is ambiguous and must be interpreted in context. The second point is that this disambiguation process is normally so efficient and so accurate that we are unaware that it is occurring. On those rare occasions when context does not immediately resolve an ambiguity, we may be amused by our error.

Figure 1 illustrates the critical role that context plays in visual perception. The figure is intended to depict the view of a train engineer who must decide whether the obstacles on the track could derail the train. The engineer would likely judge that the first animal is tiny and poses no threat, but the second animal is huge and could cause a dangerous collision. The size judgments of the engineer must be based entirely on contextual information because the pictures of the two animals are identical. The converging tracks and the gradually shrinking railroad ties provide strong depth cues that indicate that the second animal is much further away than the first animal. Our visual system uses this distance information to rescale the animals' images so that we perceive their true physical sizes. If Figure 1 depicted a real three-dimensional scene, the contextual information provided by the depth cues would produce an accurate percept.

But Figure 1 is not a real scene; it is a composite photograph made by pasting two copies of one image onto an unrelated image. If we regard the figure this way, then the contextual information provided by the depth cues is misleading and the

Figure 1 Composite Image

Notes: This picture is a composite: The image of an animal was pasted twice onto an unrelated background. Although the animal images are identical, the top animal appears larger because the background causes it to appear further away.

differential size percept is an illusion. To see the animals as identical, we would need to isolate them from the background and see them as separated on the page, not as separated in depth. We are unable to do this because, like much of perceptual processing, this context effect is not subject to conscious control.

The size illusion in Figure 1, and our inability to overcome it, reveal the automatic and decisive role that context plays in our percepts. As the four examples in this entry illustrate, similar context effects pervade every aspect of perception, from such seemingly elemental percepts as surface lightness to such high-level processes as object recognition.

Lightness Constancy

In our everyday experience, surfaces have a constant lightness: A piece of white paper looks white whether we view it inside under dim illumination or outside in bright sunlight This perceptual constancy may give the impression that lightness is a property of the paper that can be measured directly from the light it reflects, independent of context. This impression is wrong, however, because it is relatively easy to create a situation in which a white piece of paper has a lower luminance (reflects less light) than a black piece of paper does. This can be achieved by casting a shadow over the white paper and directly illuminating the black paper, or by turning the white paper away from the light and turning the black paper toward the light. Even though these changes in illumination and surface orientation have a profound effect on the luminance of the paper, they have little effect on its perceived lightness. To achieve this constancy, our visual system uses several forms of contextual information to disentangle surface lightness from the other factors that affect luminance.

The processes that underlie lightness constancy may begin early in vision. Neurons in the eye measure the luminance levels within a small region on the retina. Each neuron responds when a spot of light illuminates the center of the region that it monitors. If the intensity of the spotlight is increased, the response of the neuron will increase, but if the area of the spotlight is increased, the response of the neuron will decrease. This is because each neuron compares the amount of light in a central region with the amount of light in the surrounding area and then generates a response that reflects the difference. Because the neurons calculate luminance differences, or contrast, their responses are largely unaffected by whether the region has a uniformly high luminance level or a uniformly low luminance level. This property makes the neurons less sensitive to the effects of illumination (i.e., whether the region is brightly lit or dimly lit) because illumination tends to be uniform over small regions. By responding most strongly to a patch of light only if it differs from its surround, these neurons use local context information to begin the recovery of surface lightness. Full lightness constancy involves more complex calculations that compare luminance levels over larger areas, but all of these calculations rely on contextual information.

Motion Detection

We are extremely sensitive to motion. Even a small motion in peripheral vision can immediately summon attention. It may seem unremarkable that a moving object would stand out in an otherwise stationary world, but the image of the world on our retina is far from stationary. Because of the movements of our bodies, our heads, and our eyes, the retinal image changes constantly and dramatically. To perceive an object as moving we must disentangle the retinal motion it produces from the retinal motions that we produce by our own movements.

The process of disentangling object motions from self-produced motions is possible because self-produced motions generate distinctive and characteristic retinal flow-fields: An eye movement to the right produces a full-field translation to the left; walking forward produces an expansion pattern that radiates outward. The visual system attributes these distinctive patterns to self-motion because it is unlikely that object motions would produce them. And when object motions do mimic these patterns, they can create an illusion of self-motion. For example, drivers sometimes experience illusory self-motion when waiting at a stoplight. When the other cars start to move forward, the motion they create in the driver's periphery can cause the illusion that the driver's car is rolling backward. In addition to the information provided by flow patterns, the visual system also receives information about self-movement from the vestibular system and from neurons that control and monitor eye and neck muscles. For example, when a motor signal is sent to the eye muscles to produce an eye movement, a copy of the signal is sent to the visual system. The visual system uses these extra-retinal signals to anticipate the retinal flow pattern that the self-motion will produce.

The process of disentangling object motion from self-motion likely involves neurons in the medial temporal (MT) area of the brain. Although this brain region is located several processing stages beyond the eye, many of its neurons have a similar type of center-surround organization as the eye's contrast sensitive neurons do. The neurons in area MT monitor a particular spatial location in the visual field and respond whenever a motion with a specific speed and direction occurs at that location. When the same motion is present in the surrounding area, however, the response of the neuron is suppressed. As a result of this surround suppression, these neurons respond weakly to the type of full-field motion produced by an eye movement. For many MT neurons, surround motion in the opposite direction produces an enhanced response, making these neurons especially sensitive to the motion gradients that arise when an object moves relative to its background. The local motion differences that are calculated in area MT are sent to neighboring brain regions that use more extensive contextual information, including extra-retinal signals, to fully segregate object-motions from self-motion. Only after analyzing the complete pattern of retinal motions and the extra-retinal signals can the visual system arrive at an unambiguous interpretation of a moving stimulus.

Object Recognition

Most people have had the frustrating experience of being unable to place a familiar face. This failure of face recognition seems to occur most frequently when we know someone from one context and then meet him or her in a different context. Context may play a similar role in the recognition of all objects, not just faces. Through experience, the visual system may learn that certain objects typically occur in certain locations in certain kinds of scenes. This knowledge can facilitate recognition, especially if the stimulus is ambiguous: On a desk, a small pink rectangle might be recognized as an eraser, but on a bathroom counter, the same stimulus might be recognized as a bar of soap. Because perceptual processes are largely subconscious, our perceptual knowledge of contextual associations may be independent of our conscious knowledge of these associations. Our perceptual system can exhibit knowledge that we do not know we know.

Object recognition may be facilitated by contextual information that pertains to a particular familiar location or to a general scene category. Our facility with this second type of contextual information allows us to interpret the sequence of unrelated images that we encounter when flipping through a magazine or surfing television channels. Within a fraction of a second, we can recognize whether an image depicts a city, a beach, a desert, or a kitchen. It might be assumed that we classify scenes based on the objects they contain, but recent research suggests that the opposite is true. When classifying an image as a cityscape, we do not first recognize buildings, cars, or street signs; instead, we simply detect the predominance of vertically oriented lines. Once a scene is classified as a cityscape, this contextual information may prime us to recognize buildings, cars, signs, and other objects that are likely to appear in this context.

The brain processes that underlie object recognition and scene categorization are not fully understood, and so the characterization of their interaction is still speculative. Unlike the previously described processes, which involved comparisons of similar information within a brain region, this context effect likely involves interactions between two brain regions that process different types of information: the parahippocampal area (PHA), which processes scene information, and the lateral occipital complex (LOC), which processes object information. For scene categorization to influence object recognition, the PHA would presumably relay the results of its processing to the LOC. These results would then bias processing in the LOC to favor objects associated with the scene category.

Phoneme Recognition

The Groucho Marx quote at the beginning of this entry demonstrated one role that context plays in speech perception: When a word has multiple meanings, we use context to select the meaning intended by the speaker. But context is involved at an even earlier level in speech perception because it allows us to recognize sounds as words. We perceive spoken words as a sequence of vowel and consonant sounds, that is, as a sequence of phonemes. Perceptually, each phoneme seems fixed, but acoustically, a phoneme can vary considerably depending on context. We hear the same /d/ sound in *die* and *due,* but the initial auditory signal of these words is different. This difference arises because of the flowing nature of speech production. As we articulate one phoneme, we are preparing to articulate the next, and the resulting co-articulation is reflected in the acoustical signal. The processes that underlie speech perception use contextual information to automatically and effortlessly compensate for the effect of co-articulation, resulting in a phonemic constancy similar to the lightness constancy described earlier.

In some situations, such as a noisy party or a crowded restaurant, the acoustical signal may be insufficient for reliable speech recognition. In these situations, we may rely heavily on lipreading to understand what is being said. But even when the speech signal is clear and we are not consciously paying attention to the speaker's mouth, visual context can influence phoneme perception. This automatic processing of facial movements explains why a dubbed movie can be particularly difficult to understand: We cannot ignore the visual context even when it is misleading.

Like the other aspects of perception described here, phoneme recognition is a topic of intensive research. Although this process is not completely understood, a full description will involve the integration of contextual information across time and across senses. Although it may seem paradoxical, we perceive phonemes as having stable properties that are independent of context precisely because we are so adept at interpreting them in context.

Mary Bravo

See also Constancy; Lightness Constancy; Motion Perception; Nonveridical Perception; Speech Perception; Vision; Visual Scene Perception

Further Readings

Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 339–351). Cambridge: MIT Press.

Bar, M. (2004). Visual objects in context. *Nature Reviews: Neuroscience, 5,* 617–629.

Britten, K. H. (2008). Mechanisms of self-motion perception. *Annual Review of Neuroscience, 31,* 389–410.

Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech perception. *Annual Review of Psychology, 55,* 149–179.

Recanzone, G. H., & Sutter, M. L. (2008). The biological basis of audition. *Annual Review of Psychology, 59,* 119–142.

Schwartz, O., Hsu, A., & Dayan, P. (2007). Space and time in visual context. *Nature Reviews Neuroscience, 8,* 522–535.

Shevell, S. K., & Kingdom, F. A. A. (2008). Color in complex scenes. *Annual Review of Psychology, 59,* 143–166.

CONTRAST ENHANCEMENT AT BORDERS

Nineteenth-century visual scientists such as Michel-Eugène Chevreul, Ernst Mach, Hermann von Helmholtz, Ewald Hering, and Johannes Peter Müller discovered that simultaneously presented stimuli could affect each other's perceived contrast. For example, notice how each of the solid stripes in Figure 1 appears lighter on the left than on the right, even though each stripe has the same physical intensity across its width. This illusion is called "Mach bands," and it illustrates how the contrast of a stimulus is enhanced at its borders. This entry describes spatial and temporal contrast enhancement at borders.

Spatial Contrast Enhancement at Borders

Visual spatial contrast is the perceived difference in brightness or color between two or more locations in the visual scene. Perceived contrast is determined by the physical difference in intensity and color between two adjoining areas. Thus, contrast is not

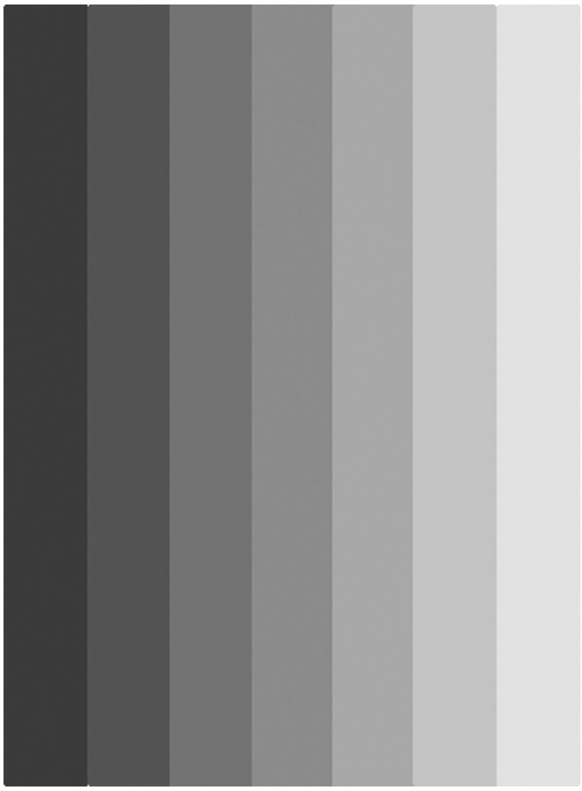


Figure 1 Mach Bands

Source: Chevreul, M. E. (1987).

Notes: This Mach band demonstration was originally designed by Chevreul in 1839. Notice how each vertical stripe appears to be lighter on the left than on the right. This illusory effect is caused by contrast enhancement at the borders.

a physical quantity but a perceptual comparison between two areas. The neural computation underlying the perception of contrast is partly carried out by lateral inhibition circuits in the visual system. Lateral inhibition is the process by which an excited neuron suppresses the activity of its neighbors across visual space. Lateral inhibition circuits are found in all levels of the visual system (such as in the bipolar and ganglion cells of the retina) as well as in other sensory systems. One perceptual consequence of lateral inhibition is that stimuli to both sides of a luminance border are differentially enhanced in an illusory fashion (as in Figure 1).

Another corollary of contrast enhancement at borders is that non-borders (such as the interiors of objects) are perceptually suppressed by the same lateral inhibition circuits that enhance the borders. Indeed, neurons in early visual areas are activated only by the borders (edges or corners) of stimuli, not by their interiors. The process is metabolically efficient in that it reduces the metabolic demand required to respond to the presentation of a stimulus. Only the neurons at the edges of surfaces require energy to respond, whereas the neurons in the interiors do not require as much energy (because they are quiescent). Thus, our visual system sees the edges of surfaces, and then uses this information to create the middles, through a process called filling-in. The neural circuits underlying filling-in have been localized to cortical visual areas beyond area V1 (such as areas V2 and V3).

Temporal Contrast Enhancement at Borders

Contrast is increased at their temporal borders (that is, their onsets and terminations) as well as at the spatial borders of stimuli, also because of lateral inhibition. And just as the interiors of spatial stimuli are suppressed when compared with their borders, so too are the temporal mid-lives of stimuli suppressed when compared with their onset and termination. The result is that the perceived contrast of a stimulus is highest just after it turns on and then again after it turns off. Visual masking (the effect in which the visibility of a target stimulus is reduced by a masking stimulus that does not overlap the target in space or time) occurs perceptually when the neural responses to the target onset or termination are inhibited.

Stephen L. Macknik and Susana Martinez-Conde

See also Contrast Perception; Lateral Inhibition; Light Measurement; Receptive Fields; Vision: Temporal Factors; Visual Illusions

Further Readings

Chevreul, M. E. (1987). *The principles of harmony and contrast of colors and their applications to the arts* (C. Martel, Trans.). West Chester, PA: Schiffer.

Macknik, S. L., & Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nature Neuroscience, 1*(2), 144–149.

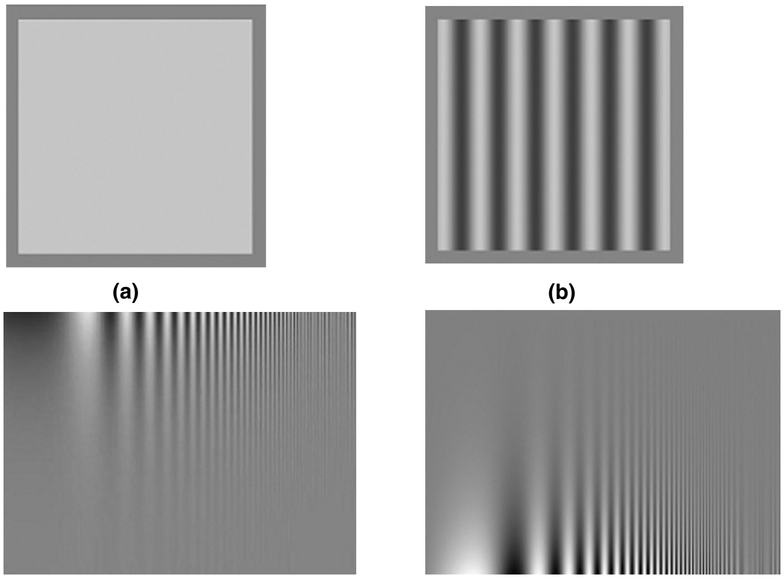
Macknik, S. L., Martinez-Conde, S., & Haglund, M. M. (2000). The role of spatiotemporal edges in visibility and visual masking. *Proceedings of the National Academy of Science USA, 97*(13), 7556–7560.

Troncoso, X. G., Macknik, S. L., & Martinez-Conde, S. (2005). Novel visual illusions related to Vasarely's "nested squares" show that corner salience varies with corner angle. *Perception, 34*(4), 409–420.

CONTRAST PERCEPTION

Sensitivity to light is common in the animal word. Even the one-celled euglena have eyespots that enable them to sense light. Only animals with eyes can sense differences in the light coming from different objects in their environment (*contrast*). Humans sense the intensity of light, but they are much more sensitive to contrast. What this means is that when intensities change slowly, as at sunset, we can tell that the intensity of light is changing, but we have a much better sense of differences in brightness between one object and another (*contrast perception*). Sensitivity to contrast makes it possible for us to see in detail what is around us and to do so over a huge range of intensities. This ability depends critically on the fact that different objects reflect light differently. Both the intensity and the spectrum (wavelengths) of light vary from one region of a scene to another. For example, in a scene with grass and a dark gray road, the grass will reflect more light and the light will be concentrated in the middle wavelengths; the road will reflect less light and it will have a broad spectrum of wavelengths. In this example, as in most real-world situations, both brightness contrast and color contrast are present. This entry focuses on brightness contrast perception.

There are several different measures of contrast, but they all indicate the *ratio* of light intensity in one location to the intensity in another. For scenes that are seen by reflected light, as most scenes are, the contrast remains constant as the light falling on the scene (illumination) changes. For example, going from bright daylight to dusk, the contrast (ratio) between the intensities of the grass and the road remains the same. Many visual systems, including ours, are designed to be sensitive to contrast.



(a) (b) (c) (d)

Figure 1 Spatial Patterns and Contrast Perception

Notes: (a) A uniform gray rectangle on a darker background. (b) A sine wave grating with the same contrast as the rectangle and background in 1(a). (c) A grating that decreases in stripe width from left to right and decreases in contrast from bottom to top. At normal reading distance, stripes of middle widths are visible higher in the image and thus at lower contrasts than are wide or narrow stripes. We see intermediate stripe widths better than we see wide or narrow stripe widths. The stripes that are seen best will vary with distance because the width in the retinal image determines visibility. An image like this was first created by Fergus Campbell and John Robson in 1964.

If a system were completely insensitive to absolute intensity, but sensitive to ratios of intensity, it would produce a constant response to a scene as the illumination on it changes. Our visual systems are not completely insensitive to absolute intensity, but they are much more sensitive to contrast. The process that brings about this difference is *light adaptation,* which has the effect of reducing sensitivity to light as the intensity increases and increasing sensitivity as the intensity decreases, while maintaining high sensitivity to contrast. Light adaptation does this by decreasing all the intensity signals in the visual system by the same fraction, leaving the ratios among the intensity signals (contrast) relatively unchanged.

Seeing Contrast Patterns

Any variation of intensity over space is referred to as a spatial pattern. Some patterns are much more visible than others are, and these variations in sensitivity provide evidence about the neural units and processes in the brain that bring about pattern vision. Differences in the visibility of patterns are often determined by measuring the *absolute contrast threshold,* the minimum contrast at which a pattern is seen. Low thresholds correspond to high sensitivity. Small patterns of stripes, called gratings (Figure 1b), have lower thresholds than do uniform surfaces (Figure 1a) and have frequently been used to study contrast sensitivity. There are two reasons for this. First, the visual system is sensitive to these patterns, which suggests that it has units tuned to such patterns, and second, gratings have properties that are useful in developing models of pattern vision.

We might guess that we would be most sensitive to the widest stripes, but we are actually most sensitive to stripes of intermediate width. This is illustrated in Figure 1(c). Many factors contribute to this variation in sensitivity, but the principal one is that the neurons in the brain that play the central role in the detection of patterns are most sensitive to gratings with intermediate stripe widths. These are called pattern vision units, even though higher-level units are also involved in pattern vision. For each of these neural units, there is a region of the retina (receptive field) such that light in that region will affect the response of the unit. The effect is either to increase the response of the unit (excitation) or to decrease its response (inhibition). The receptive fields of many cortical units are divided into striped regions that are alternately excitatory and inhibitory. The threshold for a particular pattern depends on how closely it matches the receptive fields of the detecting units. If the correspondence is close, the response will be large and less contrast will be needed to reach threshold. If the stripes of the pattern are made either wider or narrower, the response will be less. Evidence from single unit recording indicates that many neurons in animal brains have striped receptive fields (simple cells).

Seeing Differences in Contrast

Although much research on contrast perception has focused on absolute thresholds, the *contrast difference threshold* provides important insights into the process of contrast perception over the whole range of contrasts. To measure the difference threshold for contrast, a pattern is selected and set at the contrast at which the threshold is to be measured (base contrast). A small increase is added to the base contrast to produce a pattern of higher contrast. The two patterns are presented in random order and the observer indicates which pattern has the higher contrast. The added contrast is varied to determine the contrast difference threshold (minimum contrast difference that can be seen). As the base contrast increases from 0, the difference threshold at first decreases and then increases. In other words, we see differences in contrast best neither at absolute threshold nor at high contrast, but at an intermediate contrast. This result is interpreted as indicating that as the base contrast increases from 0, the response of contrast sensitive neurons in the perceptual system increases slowly, then rapidly, and then slowly. This has been confirmed by directly recording from V1 neurons in animals.

The Neural Code for Contrast Patterns

The evidence presented so far is consistent with the pattern vision units all having the same type of receptive field, differing only in their positions on the retina. However, there is both psychophysical and biological evidence for the existence of units tuned to different stripe widths and orientations. These units are in the visual cortex of the brain, not the eye, and they receive input from many

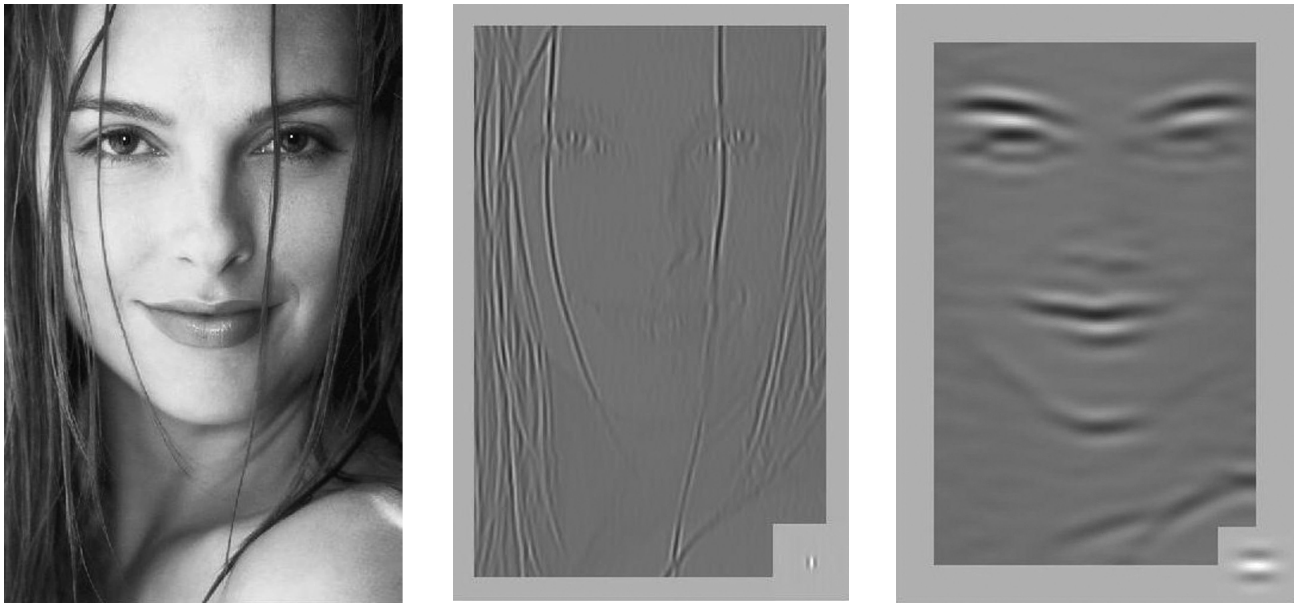


Figure 2 The Effect of Filtering a Natural Image

Source: © 2009 Jupiterimages Corporation.

Notes: (a) Gray-level image. (b) Image filtered with a vertical filter with narrow stripes. (c) Image filtered with a horizontal filter with broad stripes. The two filtered images represent two of the neural images in the multi-neural-image representation. The small images in the lower right represent the receptive fields. Bright regions correspond to excitation and dark regions to inhibition.

photoreceptors in the eye. Much of the psychophysical evidence comes from studies of *selective adaptation* and *masking* in human vision. If one pattern is viewed for several seconds before a second is presented, the first pattern will produce a decrease in sensitivity to the second (increase its threshold), but this happens only if the patterns are similar (selective adaptation). Likewise, when one pattern is superimposed on another, the sensitivity to that other pattern will decrease (masking) only if the patterns are similar in stripe width, orientation, and position. Models of contrast perception have been based on such evidence. A key assumption of these models is that a context pattern can affect the visibility of a test pattern only if the unit that detects the test pattern is either excited or inhibited by the context pattern. Thus, if a pattern is seen, but has no effect on the visibility of a second pattern, different neural units mediate the perception of the two patterns and the effects of the patterns are independent. If a pattern affects the visibility of the second pattern, the first pattern must influence the response of the units that detect the second pattern. Single unit recording has shown that neurons in area V1 of the cortex differ in the stripe widths and orientations, as well as the regions of the retina that they are most sensitive to.

Models of Contrast Perception

Current models of the contrast perception system assume an array of neurons with receptive fields tuned to different stripe widths, orientations, and positions. Each neuron sums excitation and inhibition over its receptive field and its response is an S-shaped function of this sum. Each small region of retina has a set of overlapping receptive fields tuned to different stripe widths and orientations. These arrays of receptive fields are repeated over the retina, but we do not know in detail what the sensitivities of these receptive fields are or how they are distributed. It would be sufficient for vision to have units tuned to a few stripe widths and orientations, and models often assume this, but single unit recording has found units tuned to

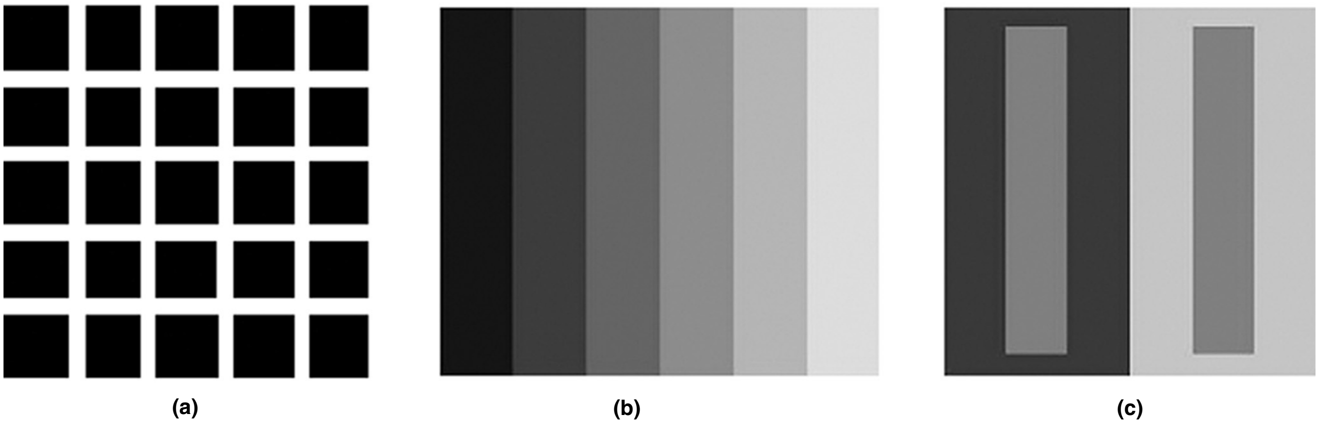


Figure 3 Three Contrast Illusions

Notes: (a) Hermann grid. The dark spots seen at the intersections of the white lines are not present in the stimulus. (b) Mach bands. Each step is uniform in intensity. That a bright line is seen to the right of each vertical edge and a dark line is to the left are created by the visual system. (c) Simultaneous brightness contrast. The two center rectangles are equal in intensity. The one on the left appears lighter. In each case, the region surrounding the region where the illusory percept appears is critical to producing the illusion. If you cover the surrounding regions, the illusion will disappear.

many spatial frequencies and orientations. The responses of these neurons together contain much of the information in the image, but the information is in a neural code. Imagine a set of neurons whose receptive fields lie in different places on the retina, but all respond to the same stripe width and orientation. The pattern of responses over this set of neurons may be thought of as a neural image that contains only those features of the retinal image that have the corresponding stripe width and orientation. Other sets of neurons have patterns of activation that represent features of other stripe widths and orientations. Thus, the image on the retina is represented by many different neural images, each containing part of the information needed to reconstruct the original image. For example, a set of horizontally oriented receptive fields with wide stripe widths produces a neural image like that in Figure 2(b). This is a blurry version of the original in Figure 2(a) emphasizing horizontal contours. The neural image produced by a set of vertical receptive fields with narrow stripe widths codes for fine detail that is vertically oriented (Figure 2c).

Two critical unanswered questions in contrast perception are these: What is the function of this multi-neural-image representation, and what role does it play in the generation of percepts? One function may be to represent the image in a way that minimizes the number of bits of information required to code natural images (a representationally efficient code). Although a huge variety of natural images are possible, natural images tend to have many properties in common. For example, in natural images, the contrasts at neighboring points tend to be similar, except at edges. Thus, natural images form only a relatively small subset of all possible images. Evidence indicates that the multi-neural-image code used by the visual system efficiently encodes natural images.

Consider what happens when the image of a complex scene lands on the retina. A set of neural images will be formed. Figures 2(b) and 2(c) represent two such images, but there would be many more. Contrast will vary over space in a different way in each of them. Objects are delineated by edges, so it is essential to find the edges in a scene. Where there is an edge between regions of different reflectance, there will be a change in contrast in those neural images with the same orientation as the edge. If such a change always indicated an edge, finding edges would be easy. However, differences in lighting, orientation of surfaces, shadows, and texture within a region will also produce changes in contrast. To perceive what is in the scene, the visual system has to disambiguate and interpret these changes in contrast. There are several models of how a multi-neural-image

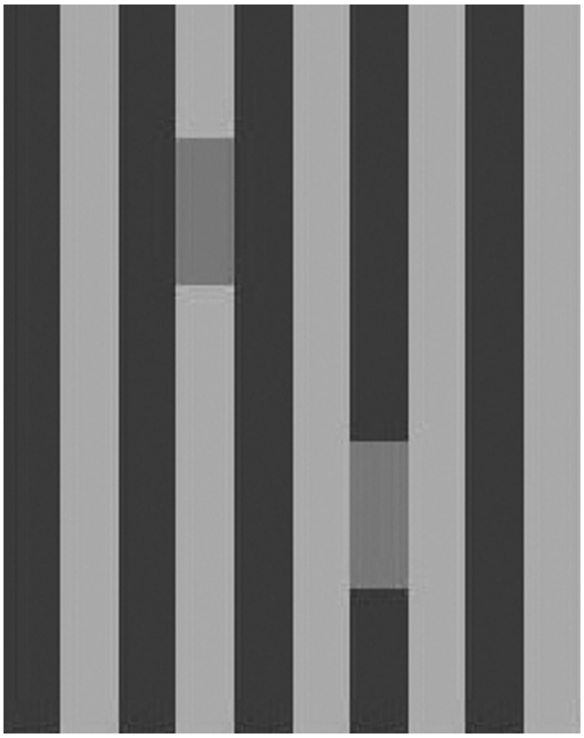


Figure 4 White's Illusion

Note: The rectangle in the dark stripe appears brighter than the rectangle in the bright stripe even though the first has intense regions on either side and the second has dark regions on either side.

representation might be used to do this. These analyze and combine the different neural images to extract edges, lines, bumps, and other elementary features in the scene.

Context Effects in Contrast Perception

The visibility and perceived contrast of a test pattern are greatly affected by other patterns that are near it in space and time; they need not be superimposed. These *context effects* present a challenge to models of contrast perception. The predominant effect of context patterns (*masking*) is to make the pattern harder to see or to reduce perceived contrast. Masking is thought to occur when the context patterns inhibit the unit that detects the test pattern. In some cases, context patterns make other patterns easier to see and higher in perceived contrast (*facilitation*). Low-contrast patterns that are similar to the test pattern often facilitate perception, whether they are superimposed over the pattern or adjacent to it. Facilitation is produced when the context pattern excites the unit that detects the test pattern. Masking and facilitation effects are complex and are not fully understood.

Illusions in Contrast Perception

An important fact about contrast perception is that our experience does not always correspond to the contrast pattern of the stimulus. Another way to say this is that there are many illusions in contrast perception. Three of the most familiar contrast illusions are the spots seen at the intersections of the Hermann grid, the dark and bright regions seen adjacent to edges (Mach bands), and the lower brightness of a region when it is surrounded by a region of high intensity (simultaneous brightness contrast). See Figure 3(a, b, c). There are many other contrast illusions. Models of these illusions are often based on units with concentric or striped receptive fields together with the assumption that the brightness at a point corresponds to the response of the receptive field centered at that point. This simple idea is central to explanations of these phenomena. For example, in Figure 3(c) imagine that there is a receptive field with an excitatory region centered on each of the center stripes and inhibitory regions on either side. Both receptive fields will be excited equally, but the one on the right will be inhibited more, and therefore this cell will respond less and we will perceive the center stripe on the right to be less bright. There are many contrast illusions that this simple model does not explain. White's illusion (Figure 4) is an example. Here the rectangle with more intense light on either side appears brighter. Some models combine the responses of many V1 type cells to explain these phenomena, but cognitive organizational processes at higher levels in the visual system explain the phenomena in other models. No current model describes all the phenomena of contrast perception.

John M. Foley

See also Constancy; Context Effects in Perception; Contrast Enhancement at Borders; Lightness Constancy; Neural Representation/Coding; Psychophysical Approach; Receptive Fields; Visual Acuity; Visual Illusions

Further Readings

Blakeslee, B., & McCourt, M. E. (2004). A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. *Vision Research, 44,* 2483–2503.

Bruce, V., Green, P. R., & Georgeson, M. A. (1996). *Visual perception: Physiology, psychology, and ecology.* East Sussex, UK: Psychology Press.

Graham, D. J., & Field, D. J. (2008). Efficient coding of natural images. In L. R. Squire (Ed.), *Encyclopedia of neuroscience.* San Diego, CA: Elsevier.

Graham, N. V. S. (1989). *Visual pattern analyzers.* New York: Oxford University Press.

Regan, D. (2000). *Human perception of objects: Early visual processing of spatial form defined by luminance, color, texture, motion, and binocular disparity.* Sunderland, MA: Sinauer Associates.

COROLLARY DISCHARGE

Imagine an animal walking through a forest. As it moves, branches brush against its skin, twigs snap at its feet, and patterns of light and shade alternate across its eyes. In principle, the animal should be startled by these sensory events. The activation of its skin receptors could be interpreted as caused by a fly landing on its leg, and the sounds and shadows as meaning a predator is looming. Surprisingly, the animal is not startled by these sensory events; they are expected, partly because the animal has access to an internal report of its own movements called *corollary discharge.*

Each of an animal's movements is initiated by motor commands originating from movement areas of the brain that travel peripherally to activate the appropriate muscles. Neural copies of the movement commands are issued simultaneously and travel in the opposite direction, impinging on sensory brain areas. These corollary discharge signals inform the sensory areas of the upcoming movements and allow them to prepare for the sensory consequences of the movement. As a result, our animal in the forest is not surprised by the brush of the branch, the snap of the twig, or the change in shade. Were the animal at rest, or moving without the benefit of corollary discharge, the sensory events would be startling indeed.

As a theoretical concept, corollary discharge has a rich history. Behavioral and psychophysical evidence for corollary discharge has been around for more than a century, and direct physiological evidence has accrued within the past few decades. Corollary discharge circuits have been found, for example, that transiently inhibit a cricket's hearing while it chirps. Researchers now know corollary discharge to be ubiquitous. It is present in virtually every animal species and coordinates nearly every motor and sensory system. The most is known about corollary discharge circuits that interact with the senses of vision, audition, and somatosensation, as described in this entry.

Vision

Well-understood corollary systems reside in the visual and eye movement networks of vertebrates, particularly in primates. A common visuomotor behavior of the primate that requires corollary discharge is the saccade, a fast eye movement that occurs about 2 to 3 times per second as an animal scans a scene. Saccades are beneficial because they permit rapid relocation of the fovea, but are costly because each saccade results in two unfortunate consequences for the visual system: an intrasaccadic smearing of the image and a trans-saccadic displacement of the image across the retina. With each saccade, the world should appear as both blurry and jumpy because this is what the retina is actually reporting. Instead, the world appears focused and stable despite the retinal report because the visual system receives advance warning of each saccade in the form of a corollary discharge.

Corollary discharges of saccadic eye movements emerge from oculomotor structures that span virtually all levels of the nervous system and impinge on visual areas throughout the brain, early and late, subcortical and cortical. At their point of termination, the corollary discharges interact with recipient sensory areas to minimize the effects of the eye movement–induced sensory inputs. For the case of retinal blur, the corollary discharges participate in a mechanism known as saccadic suppression. Neurons participating in saccadic suppression are transiently inhibited by corollary discharges at the time of the saccade, thus reducing the amount of visual information they convey and ultimately the amount of blur that is perceived. As for the displacement problem, one putative compensatory mechanism involves neurons that shift their receptive fields before each eye movement. Instructed by corollary discharge, these neurons sample a new part of the visual field before the eyes begin to move. By sampling the same portion of space both before and after the eye movement, shifting receptive fields effectively seem to test whether the external world moves during the saccade. If the presaccadic and postsaccadic samples match, the world is judged as stable. The corollary discharge that shifts receptive fields, at least those of neurons in the frontal cortex, is known to arise from the midbrain superior colliculus.

Audition

Corollary discharges are also important for coordinating the vocal and auditory systems. One problem faced by many vocalizing animals is the effect their high intensity voices have on their hearing. This problem can be considerable for animals that vocalize at high decibel level. Such animals as the cricket generate high intensity sounds that should overwhelm their auditory circuits, rendering them temporarily deaf to independently incoming sounds. This scenario is avoided by auditory-vocal corollary discharge mechanisms analogous to saccadic suppression that enable tight coordination between auditory and vocal systems. With each vocalization, the corollary discharges exert an inhibitory influence on neurons of the auditory system. This prevents desensitization and ensures maximal attunement to external auditory events.

In humans, corollary discharge mechanisms involving speech production and auditory regions are thought to play a role in distinguishing self-generated speech sounds from other-generated speech sounds. Neurons in language areas of the temporal lobes are suppressed during self-generated speech, but remain fully responsive when others are talking. Interestingly, in schizophrenia, the auditory dampening effect is reduced, betraying a reduction in speech production/auditory coordination. An alteration in this coordination may result in internally generated events such as thoughts being experienced as having an external source. This may serve as a basis for certain aspects of the disease pathology such as auditory hallucinations and self-monitoring deficits.

Somatosensation

Movements of the body can result in sensory input (such as when an animal touches something) that is indistinguishable from the input caused by external agents (when something touches the animal). Corollary discharges from the skeletomotor system are critical for allowing the somatosensory system to tell the difference. They play a rather diverse role because they may amplify or suppress the responses of somatosensory neurons. The difference seems to depend on motor context. A rat exploring an object with its whiskers or a person reading Braille with her fingertips depends on corollary discharges that amplify the sensory information for detailed analysis. This contrasts with suppressive corollary discharges of the somatosensory system. These inputs, originating from premotor circuits, inhibit somatosensory neurons during voluntary movements. In effect, they provide a prediction of the sensory consequences of the movement. If the actual sensory feedback matches the predicted sensory feedback, the signals cancel and somatosensory input is not perceived. Unexpected sensory inputs, however, pass unhindered and result in a surprising somatosensory experience. An everyday example is the act of tickling. We cannot tickle ourselves because we cannot surprise ourselves. Our finger movements cause somatosensory inputs, but corollary discharges of our finger movements cancel them. If someone else moves their fingers across our skin, however, we experience the familiar “tickly” sensations at full force.

Marc A. Sommer and Trinity B. Crapse

See also Action and Vision; Animal Eye Movements; Audition; Body Perception; Haptics

Further Readings

Blakemore, S. J., Wolpert, D., & Frith, C. (2000). Why can’t you tickle yourself? *Neuroreport 11,* R11–16.

Crapse, T. B., & Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nature Reviews Neuroscience, 9*(8), 587–600.

Ford, J. M., Gray, M., Faustman, W. O., Roach, B. J., & Mathalon, D. H. (2007). Dissecting corollary discharge dysfunction in schizophrenia. *Psychophysiology, 44,* 522–529.

Grüsser, O. (1995). On the history of ideas of efference copy and reafference. *Clio Medica, 33,* 35–55.

Sommer, M. A., & Wurtz, R. H. (2008). Brain circuits for the internal monitoring of movements. *Annual Review of Neuroscience, 31,* 317–338.

CORTICAL ORGANIZATION

When neuroscientists use the term *cortex,* they usually refer to the neocortex of mammals. This neocortex is part of the outer shell or bark of the forebrain that was thought to be a new structure that evolved with mammals. Now it is widely accepted that the neocortex of mammals is the same structure as the dorsal cortex of reptiles. However, the dorsal cortex is much smaller and simpler in laminar organization than the neocortex is. Thus, the neocortex is new, in the sense that all mammals have a more complexly laminated and functionally subdivided cortex than do reptiles. Although some use the term *isocortex* for *neocortex,* the term *neocortex* is almost universally used.

When we look at the human brain, most of what we see is the neocortex (color insert, Figure 38). It occupies the outer surface of the two cerebral hemispheres of the forebrain, just behind the olfactory bulb, above the piriform or olfactory cortex, and in front of the cerebellum and brainstem (color insert, Figure 38). All mammals have a neocortex. However, the structure is highly varied in size, structural differentiation, subdivisions of functional significance, and cellular and regional specializations across mammalian species. This variability is one of the reasons why the neocortex is such an important part of the brain; the neocortex has specialized features for the various needs of different species. Differences in cortical size (Figure 1) and organization largely account for the considerable differences in the abilities and behaviors of mice and men, or even monkeys and apes. But, as this entry describes, basic features of the neocortex are also shared by most present-day mammals because they were retained from an early mammalian ancestor.

Laminar Organization of the Neocortex

The laminar organization of the neocortex is similar in most mammals in that it is a thick structure that is divided from outer surface to inner white matter into a number of layers and sublayers. Traditionally, we count six layers (Figure 2; see also color insert, Figure 39), each with a specialized role in the functional organization of the cortex. This is in contrast to the dorsal cortex of reptiles where a thin row of neurons receives inputs from other parts of the brain and project to subcortical structures.

The middle layer of the neocortex, layer 4, is the main receiving layer. Neurons in this

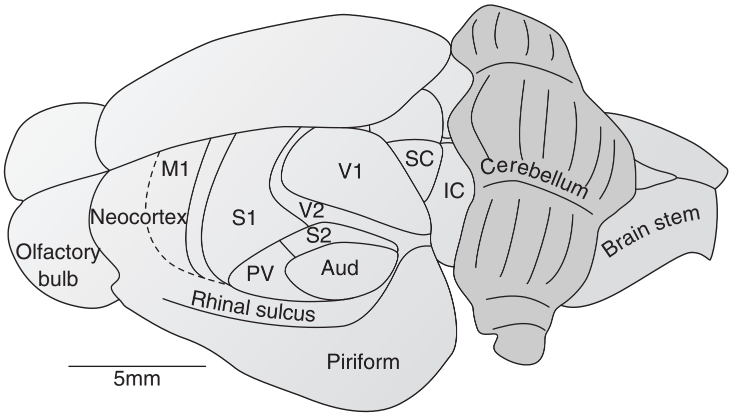


Figure 1 The Brain of a Hedgehog

Notes: The hedgehog is a mammal with a small brain. The neocortex forms the outer shell of the two cerebral hemispheres just behind the olfactory bulbs at the front of the brain (left), above the piriform or olfactory cortex below the rhinal sulcus, and in front of the more caudal superior colliculus (SC) and inferior colliculus (IC) of the midbrain, the cerebellum, and the brainstem. The neocortex is divided into a number of cortical areas including primary and secondary visual (V1 and V2) and somatosensory areas (S1 and S2), an auditory region (Aud), a parietal ventral somatosensory area (PV), and a primary motor area (M1).

layer are morphologically specialized to be activated by a small number of inputs. Thus, layer 4 neurons are small and star shaped in that short dendrites radiate out in all directions. These neurons, called granule cells because of their small size or stellate cells because of their star shape, are activated by a few axons from the thalamus, or a few axons from another cortical area. Thus, layer 4 neurons preserve the information from a few input neurons rather than integrate signals from a large number of inputs. Most of the neurons in other layers of the neocortex are pyramidal neurons, named for their pyramid shape with a long dendrite streaming from the apex toward the cortical surface. The apical dendrites typically extend into layer 1 where they are contacted by inputs with axons running parallel to the surface. Layer 4 neurons distribute information selectively to neurons above and below layer 4. Overall, small clusters of neurons in layer 4 connect with those neurons vertically aligned with them in layers above and beneath them. Neurons in other layers also tend to connect within a narrow vertical column of roughly 100 to 200 neurons. Adjacent clusters of layer 4 neurons connect in the same manner, so that the neocortex consists of arrays of narrow vertical columns of neurons extending across all layers that are directly or indirectly activated by the same small number of input neurons. Neurons above and below layer 4 are locally connected in ways that allow local circuits to derive different categories of information from the inputs, and send these different types of information to other subdivisions of the cortex as subcortical targets.

To briefly characterize the other layers of the cortex, from surface to white matter, the superficial layer 1 consists mainly of axons connecting with dendrites of cortical neurons in deeper layers. These axons largely come from brainstem neurons that are parts of modulating systems that connect with large numbers of cortical neurons, usually to enhance their responses to their dominant inputs. These modulating inputs play an important role in attention and learning. Layer 2, a narrow layer, consists of smaller neurons that make contacts with nearby neurons and the dendrites of neurons in lower layers. Layer 3 is a broader layer, especially in mammals with large brains and a lot of neocortex. Neurons in layer 3 analyze information received from layer 4 neurons, and project to other areas of the cortex. Different sublayers of layer 3 may project to different cortical areas, and they may send different types of information to these areas. Thus, the cortex has the capacity to send information to other subdivisions of the cortex, where it is processed further (serial processing). This reiteration or reprocessing is what gives the cortex its great computational capacity because simple processing steps repeated over and over can lead to sophisticated results. In addition, the same or somewhat different types of information can be

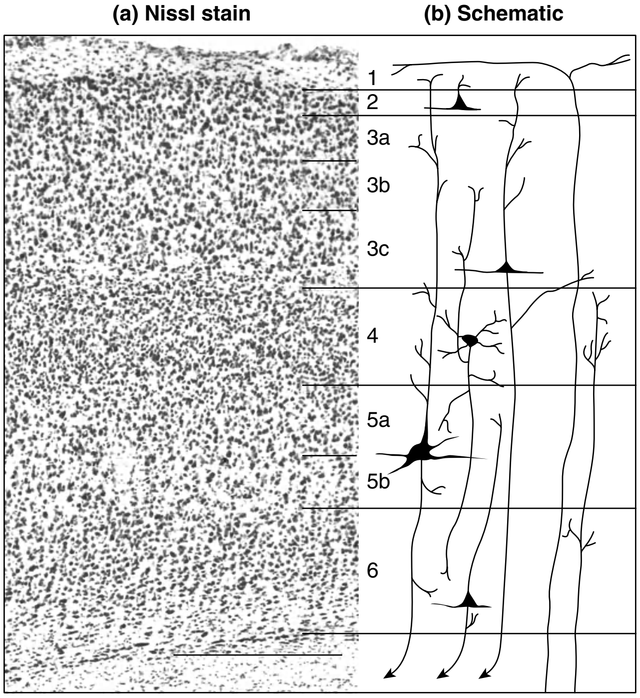


Figure 2 The Laminar Organization of the Cortex

Notes: (a) A thin section through the thickness of the cortex from the surface above to the fibers (axons that connect the cortex to other regions of the cortex and the brain) below. The Nissl stained section through V1 of a squirrel shows the locations of small, individual neurons as dark dots. Layers (1–6) are distinguished by differences in the arrangements, sizes, and shapes of neurons. (b.) A drawing of enlarged neurons (for visibility) and their local connections illustrates how vertical arrays of neurons are densely interconnected. Neurons have large cell bodies and thin dendrites and axons. Axons entering the cortex and terminating in layers 6, 4, and 1 are shown on the far right side in (b). See also color insert, figure 39.

sent from one cortical area to several others, so that information is processed in parallel streams. Thus, the cortex is characterized by both serial and parallel processing. The serial steps add processing power, and the parallel channels add processing speed.

Although most of the local connections within the cortex are vertically arranged across layers (Figure 2; see also color insert, Figure 39), some of the connections within layer 3 are horizontal and cross the vertical arrays of interconnected neurons. These horizontal connections are most dense over short distances, and few extend farther than a few millimeters. Although these horizontal connections are sparse and have weak effects, they are important in forming the complex response properties of layer 3 neurons so that these neurons are sensitive to the ongoing activities of nearby vertical columns of neurons. This is the start of a broader integration of information across larger groups of neurons in the cortex. Layer 3 neurons may also have inputs from smaller neurons in the thalamus. These inputs over their axons likely have weak but important modulating effects.

Most of layer 5 neurons are large pyramidal neurons with apical dendrites that extend up into and through layer 3. Thus, they sum much of the synaptic information of the sublayers of layer 3. These neurons project to various subcortical structures including the nuclei of the thalamus, the basal ganglia, the superior and inferior colliculi of the midbrain, and even the lower brainstem and spinal cord. For some cortical areas, a few such neurons may project to another cortical area as well. These projections help regulate ongoing behavior, influence sensory processing in other structures, and provide a further distribution of information by activating neurons in the thalamus that project back to the cortex.

The deepest layer, layer 6, has a mixed population of pyramidal and spindle-shaped cells that receive projections from upper layers of the cortex, especially layer 4, and even a small amount of thalamic input via branches of axons terminating in layer 4. These neurons, in turn, project back to the thalamic nucleus or cortical area that provides the activating input to the cortical field. Thus, layer 6 provides the important function of sending feedback to the neurons that provided the dominant activating input from the same region of the cortex. Feedback is important in modulating and adjusting the inputs to cortical neurons.

Thus, the neocortex of mammals has a laminar organization where neurons are specialized for receiving and distributing information to different targets. Species vary in the distinctness of the layers and sublayers in cell types and other features. Rather indistinct, poorly laminated layers likely reflect the primitive mammalian condition, whereas various types of morphological specialization of layers reflect more recently acquired specializations of layers as they became modified for various tasks. Most of the connectivity between neurons in the cortex is columnar, and vertical rows of neurons united by these vertical connections form a functional unit that is sometimes referred to as a mini-column. Larger collections of such vertical arrays may form functionally distinct groups of neurons called macro-columns or modules that are of two or more types within a cortical field, and variable across the cortex and across species. Variability in the types of cortical columns across cortical areas and across species provides another way in which the cortex has been modified to subserve the specific needs of different species.

Organization of the Cortex at the Local Circuit Level

The basic computations mediated by the cortex depend on the interaction of nearby neurons within layers and modules, the local circuits, as well as on the features of laminar and vertical organization. All layers of the cortex are composed of two fundamentally different types of neurons. The pyramidal neurons that project to other cortical areas or parts of the brain, and most of the layer 4 neurons discussed so far, excite other neurons and release excitatory neurotransmitters. The excitatory neurotransmitters depolarize the membrane potential of contacted neurons and thereby increase the probability that these neurons will have action potentials that stimulate other neurons. These excitatory pyramidal cells and excitatory spiny stellate and granule cells constitute roughly 80% of cortical neurons. Across all layers, the other 20% of neurons are non-spiny stellate cells and other interneurons that release an inhibitory neurotransmitter when they are excited. This transmitter hyperpolarizes synaptically contacted neurons so that they are less likely to produce action potentials (neuronal spikes). Inhibitory neurons have short axons and therefore act locally to reduce the firing rates (the production of action potentials) of pyramidal and spiny stellate neurons. Because external and intrinsic excitatory inputs tend to terminate on both excitatory and inhibitory neurons, the initial effect is excitation of excitatory neurons followed by inhibition from the inhibitory neurons. The inhibitory neurons tend to dampen the effects of persisting sources of excitation while having little effect on sudden changes in excitation. Thus, change in stimulus conditions is well detected by the cortical circuits, but less computation is devoted to persistent conditions. In addition, the inhibitory inputs often contribute to other types of computation based on temporal and spatial differences in the activations of excitatory and inhibitory neurons.

Plasticity of Local Circuits of Cortical Neurons

Another organizational feature of the cortex that makes it such a biologically useful organ is that the functions of circuits are adjustable rather than fixed. The synaptic contacts on individual neurons can be made stronger or weaker so that the activating inputs become more or less likely to cause the neuron to fire (produce action potentials) and send information to other neurons. This plasticity is possible because neurons have two types of receptors in their membranes for the excitatory neurotransmitter, glutamate. One type of receptor, the non-N-methyl-D-aspartate (non-NMDA) receptor is an excitatory channel that tends to generate neuronal spikes. If the NMDA receptor is activated when the non-NMDA receptors are activated, the activation of NMDA receptors results in internal changes in the neurons that strengthen the effectiveness of the non-NMDA synapses that were just activated. Thus, the neuron becomes more responsive to some of its inputs. This adjustment, called long-term potentiation or LTP, strengthens synapses for as long as weeks of time. Other patterns of synaptic activity weaken synapses, making some inputs less effective, a change called long-term depression or LTD. As a result, the functions of local circuits of cortical neurons are constantly adjusted so that they respond to inputs more or less forcefully. In sensory areas of the cortex, such adjustments are responsible for changes in a neuron's responsiveness to sensory stimuli. Such changes in neuron responsiveness may result in a type of performance improvement, called perceptual learning, where improvements in simple perceptual tasks follow practice.

Areas and Organs of the Cortex

The cortex is divided into a number of areas that have different functions (e.g., color insert Figure 38 and Figure 1). These areas were called the organs of the brain by an early investigator, Korbinian Brodmann, because different cortical areas are specialized for different functions, just as the heart, kidney, and liver have different functions. Although all areas of the cortex process information in basically the same way, areas differ in the inputs they receive and the brain structures they target. For example, all or nearly all mammals have a primary visual area of the cortex called V1 (visual area 1) or area 17, the 17th area described by Brodmann. V1 is distinguished from other cortical areas by having visual inputs relayed from the eye to a part of the thalamus (another part of the brain) and then to V1. As a result, neurons in V1 are activated by visual stimuli, and they participate in an initial analysis of the visual scene. The outputs of V1 go to the superior colliculus, a midbrain structure that instructs motor centers for eye movements so that the eyes are directed at objects of probable interest, such as changes in the visual scene. V1 also sends outputs to other, nearby visual areas, given various names or numbers, that process visual information further. After several such steps, the processing results in perception and the use of visual information to guide motor behavior, such as reaching for an object. Other types of sensory inputs access other cortical areas, where sensory information is analyzed in a number of steps across a sequence of sensory areas for taste, touch, or hearing. Other cortical areas combine information from areas devoted to one sense or another, to further influence perception and guide motor behavior. Still other cortical areas are involved in comparing ongoing events with past events, retaining information in short-term memory, determining the emotional significance of stimuli and events, and motivating behavior. A number of these areas access areas of the motor or sensorimotor cortex, where outputs to brainstem and spinal cord neurons that project to muscles produce movements.

Species of mammals vary in their number of cortical areas. Because areas need to be of a certain minimal size to perform their functions, mammals with small brains have few cortical areas. The least shrew, one of the smallest mammals, may have the fewest areas, on the order of 10 or so, but most small mammals have 20 to 30 areas. Macaque monkeys appear to have 80 or more, and some estimate that humans have more than 200 areas. The great advantage of having more cortical areas is that more steps are possible in the area-to-area processing sequence, leading to more sophisticated processing and more parallel processing sequences, resulting in more types of processed information. Most small-brained mammals have five or more somatosensory areas, three to five visual areas, and two to three auditory areas. Some mammals depend on the motor capacities of their somatosensory areas for cortical movement control, and placental mammals have at least two areas specialized for motor functions.

Cortical areas have been traditionally identified by differences in laminar structure as viewed in thin sections cut from the brain. Primary sensory areas, for example, are specialized for receiving sensory information from the thalamus, so that the receiving layer, layer 4, is packed with small neurons and is well developed. Motor areas have a poorly developed layer 4, and layer 4 may even be missing. Instead, motor areas have large neurons in layer 5 that project to the brainstem and spinal cord. These neurons send information a long distance via long axons, so their cell bodies in the motor cortex need to be large. Areas differ in many other ways, and the neurons within them specialize for their functional roles by expressing different proteins, receptors, and neural transmitters. But many adjoining cortical areas look much alike, and they are difficult to identify in brain sections. In addition, the same area—the primary visual cortex, for example—may look more distinct in some mammals than others because the area may be more or less specialized for its functional role.

As areas depend on differences in connections with other parts of the brain for their differing functions, areas can also be identified by their patterns of connections, which can be revealed by a number of experimental techniques. Sensory projections from the eye, ear, and skin maintain their relative spatial orders as they relay from early processing stations, brainstem nuclei, to the cortical primary sensory areas. The primary somatosensory area, S1, for example, occupies a middle portion of the cortex where it extends from the midline of the brain where the two cerebral hemispheres meet to near the lateral edge. Each hemisphere receives sensory inputs from the skin of mainly the opposite side of the body. Inputs from the toes activate neurons near the midline, whereas those from the lips activate neurons near the lateral margins of S1. Overall, the contralateral skin activates neurons from toes to lower trunk, to upper trunk, to fingers to chin, lips, and tongue in a mediolateral sequence across the cortex. Thus, S1 is said to represent the contralateral body surface. The size of the representation of different body parts varies across species as body parts vary in sensory receptor densities. Thus, humans have large representations of the fingers, which are densely packed with sensory receptors, whereas rats have a large representation of the sensory receptors of the whiskers of the upper face, which are activated when they whisk to detect objects in the dark.

Visual areas and auditory areas also represent arrays of sensory visual or auditory receptors in systematic patterns. In V1, vision of the lower, contralateral visual quadrant (upper section) is represented in anterior or medial V1, and the upper quadrant is represented in posterior or lateral V1. Humans, with a central fovea in the retina that is packed with receptors, devote a large area of V1 to these central receptors, but V1 of rats and mice is less specialized for central vision. The auditory cortex of echolocating bats is specialized for high-frequency hearing in ranges that humans cannot perceive. Thus, sensory areas vary across species in the ways they represent sensory receptors.

Cortical Modules

Areas are at least sometimes and possibly often or always divided into smaller functional divisions called cortical columns or modules. For example, the somatosensory cortex, S1, of primates, receives two types of information from the skin. One type, the slowly adapting (SA) input, is from neuron afferents that respond with neuronal spikes throughout a brief period of maintained skin touch, but the other type, rapidly adapting (RA) afferent input, responds only at the onset and offset of the touch. Remarkably, the two types of inputs are kept separate as they are relayed over two brainstem nuclei all the way to the cortex, where SA and RA inputs activate small, alternating patches of layer 4 neurons. As the information in each patch is processed in neurons in a column of cells above and below the layer 4 patches, the cortex is said to have columns of neurons that process either SA or RA information. The SA information from any skin location—the tip of the thumb, for example—is processed next to the RA information from the same skin location, resulting in two interdigitated representations of the skin of the contralateral body from toes to tongue. As another example, the primary visual cortex of humans and most primates has alternating narrow bands of layer 4 neurons that are activated by either the right or left eye. In turn, neurons in these bands activate vertical bands of the cortex, the so-called ocular dominance columns that selectively process information from the right or left eye. Other such separations of cortical areas into alternations of two or three types of processing columns or modules have been demonstrated in other cortical areas. In this way, a single cortical area can function as two or three. These types of modules within any area vary across species, so this is another way that cortical organization can vary to produce functional differences between species.

Little Cortex in Some Mammals

The neocortex emerged as a new type of structure in early mammals that differed from the thin dorsal cortex of reptiles by being a thick, laminated organ divided into areas differing in inputs and outputs. In each area, narrow, vertical arrays of 100 to 200 neurons process information from a small number of inputs, and send this information to similar vertical arrays of neurons in other cortical areas where the processing steps are repeated. This reiteration of a 100–200 neuron-processing step allows complex computations to result from the summation of a series of simple computations. Thus, the cortex can be a powerful processing organ.

The cortex is also a flexible processing organ in that it varies in organization and size as mammals have evolved and adapted to different environmental niches. Large brains with large amounts of neocortex provide a computational advantage, especially as large amounts of the cortex are likely to be subdivided into large numbers of cortical areas. The cortex specializes for the needs of different mammals by varying in how it is subdivided into areas, how areas disproportionately represent sensory inputs or motor functions, how areas are divided into modules of neurons with different functions, how areas are differentiated into layers and types of neurons, and how areas connect with other areas and nuclei. This potential for functional variability has allowed species of mammals to become different in behavior and mental capacities. In humans, and at least a number of other mammals, the neocortex, with its great information processing and information storing capacity, somehow mediates consciousness.

If the neocortex is so useful, why do some mammals have little of it? The simple answer is that large brains with large amounts of cortex take a long time to make and become fully functional. Thus, mammals with a lot of neocortex grow up slowly, and take a long time to reproduce. They cannot have short life spans. Small mammals with small brains and little neocortex (Figure 1), may have rather limited computational abilities, but they can reproduce early and often in their relatively short lives. Large brains are also metabolically costly to maintain. Mammals with large amounts of neocortex need to eat a lot of highly nutritious food. The different sizes and organizations of the neocortex, in conjunction with the rest of the brain, largely determine if you behave as a human or a mouse. Humans have large brains with a massive amount of neocortex (color insert, Figure 38). This cortex mediates complex abilities that allow humans to live a long time, protect their offspring while they develop, and obtain high-grade food that meets the metabolic demands of the large brain.

Jon H. Kaas

See also Modularity; Neural Recording; Vision

Further Readings

Kaas, J. H. (2007). Reconstructing the organization of neocortex of the first mammals and subsequent modifications. In J. H. Kaas & L. A. Krubitzer (Eds.), *Evolution of nervous systems* (Vol. 3, pp. 27–48). London: Elsevier.

Northcutt, R. G., & Kaas, J. H. (1995). The emergence and evolution of mammalian neocortex. *Trends in Neurosciences, 18,* 373–379.

Preuss, T. M. (2000). The diversity of cerebral cortex. *Brain, Behavior and Evolution, 55,* 283–347.

CORTICAL REORGANIZATION FOLLOWING DAMAGE

In humans, the cortex is a large part of the brain that is so extensive that it covers most other parts. This cortex receives inputs over brainstem and thalamic relays from the sensory receptors in the eyes, skin, ears, and other parts of the body. This information is further processed to create a model of the external world that usefully directs our behavior via the outputs of a number of motor and motor-related cortical areas. Although cortical size and organization vary greatly across the various mammalian species, all have primary sensory areas, the subdivisions of the cortex at the start of the cortical processing systems that relay to variable numbers of additional sensory, multisensory, evaluative, motivating, motor, and memory related cortical areas and subcortical structures. When damage to the nervous system impairs the functions of some part or parts of the cortex, usually the person or animal is unable to perform all of the sensorimotor, or other functions. Impairments could result from direct damage to the cortex itself, such as that following a stroke, or after damage at any level of the incoming sensory relays, such as after spinal cord damage, so that cortical areas are deprived of some of the sensory information they use to do their job.

Clinical studies of patients with damage to the nervous system have shown that impairments vary with the parts of the nervous system that are damaged, and that the impairments can be severe or mild, depending partly on the extent of the damage. But, these studies also show that considerable recoveries of lost abilities can occur, and mild impairments can completely disappear after weeks to months of time. Furthermore, some behavioral therapies seem to potentiate these spontaneous recoveries. Why do such recoveries happen, and what can be done to maximize the return of functions?

To help answer these questions one needs to know what is happening to the brain during the recovery periods. The general assumption is that the damaged nervous system is, to some extent, repairing itself. To see if this is happening, and how, it is useful to consider the results of simple, controlled experiments on animals. In cases of human stroke, the damage often involves a number of cortical areas, and the extent of the damage and the areas impaired can be highly variable across individuals. Likewise, when sensory and motor losses occur as the result of accidental spinal cord damage, many systems may be impaired, and the types and extents of damage are variable. This entry considers the results of two types of experiments. In one type, there is no direct damage to the cortex itself. Instead, a sensory loss is created by damaging a peripheral sensory nerve or another part of a sensory pathway that leads to the cortex. Thus, part of the sensory input that is used by a primary sensory area and subsequent cortical processing areas is missing. The sensory and motor impairments that follow such damage usually diminish over time, but why? The second type of experiment is one where part of a cortical area is directly damaged. If behavioral improvements follow the initial impairments after such lesions of the cortex, do they depend on compensations based on other parts of the nervous system, or on internal alterations within the damaged areas? The results of controlled experiments of both types indicate that both the loss of activating inputs and direct damage to parts of cortical areas are followed by a sequence of changes in the internal organization of the affected cortical areas, and these structural and functional reorganizations are, at least partly, responsible for some of the recoveries that do occur.

Consequence of Sensory Loss

Sensory areas of the cortex are a good place to determine the consequence of a loss of inputs because their internal organizations can be reliably determined, both before and after a sensory loss. Sensory pathways from the peripheral receptors to the cortical areas maintain topographic organizations that reflect the spatial arrangement of receptors on the receptor sheet. Thus, the afferents from touch receptors in the skin of the hind limb, trunk, forelimb, and face project to a nuclear complex in the lower brainstem in a sequence. This somatotopic order is maintained in the projection from this brainstem nuclear complex to the ventroposterior nucleus of the contralateral thalamus, which in turn projects in the same sequence of body parts to the primary somatosensory cortex (area 3b). In humans and other primates, this means that the most medial part of the somatosensory cortex is activated by touching the foot and hind limb, a more lateral part by the hand and forelimb, and a most lateral part by the face and tongue (Figure 1). The primary area of somatosensory cortex is thus said to “represent” the contralateral body surface from toes to tongue in a mediolateral sequence. Within a lateral part representing the hand, the digits 1 through 5 are presented in a sequence, next to the pads of the palm. The representation of the hand, and other parts of the body, is remarkably consistent across individuals of the same species. This organization can be revealed in several ways, most usually by recording from neurons throughout the area with microelectrodes and determining where on the body light touch activates neurons at each recording site. This is called mapping the somatosensory cortex. The organization of the map can be determined in great detail, so any changes in the map produced by a sensory loss can be detected by a mapping experiment after the recovery period.

Because of its experimental advantages, the somatosensory cortex has been the most investigated cortical area for the consequences of sensory loss. One of the first of this type of investigation was to cut a sensory nerve to the thumb side of the hand in monkeys. This produced a loss of tactile inputs from the thumb, digit 2, part of digit 3, and the adjoining palm. Immediately after such a sensory loss, the part of the contralateral somatosensory cortex that is normally activated by the missing inputs failed to respond to touch on the denervated skin or on any other part of the hand or body. However, within weeks of post-lesion recovery, the deactivated part of the hand representation became responsive to the preserved inputs from the hand, mainly

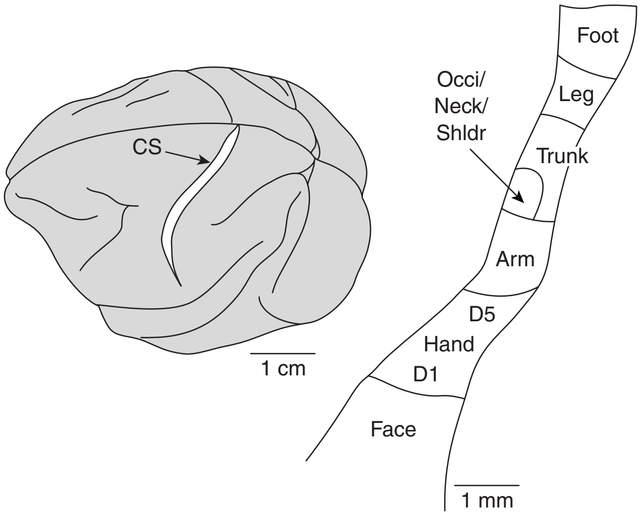


Figure 1 The Brain of a Macaque Monkey Showing the Location of the Central Sulcus (CS), in Which Most of the Primary Somatosensory Cortex (Area 3b) Is Hidden

Notes: The way touch receptors on the body surface are represented from medial to lateral in the brain is shown on the right. The motor cortex is in front (to the left) of the somatosensory cortex, and it contains a similar representation, but one of body movements.

those from the back of the digits and hand. Thus, deactivated portions of the cortex did not remain deactivated because preserved inputs took over by expanding into the cortex that originally responded to the damaged inputs, thereby activating neurons that would otherwise be unresponsive to sensory inputs. Other experiments with different types of sensory loss also produced cortical reorganization. The key result was that neurons throughout even large extents of deactivated cortex recovered responsiveness to touch, but on parts of the body surface with remaining inputs. This process became functionally significant when a few scattered inputs from the hand remained. Normally, these few intact inputs would activate few cortical neurons, or even none if the inputs were too sparse, but during weeks of recovery, these inputs come to activate large numbers of neurons throughout the deprived zone of the cortex. With this recovery of cortical activity evoked by touch on the hand, the ability to use the hand in tactile discrimination and the sensory control of motor behavior returned. The important lesson here is that a few preserved sensory inputs can do a lot. But how do they become more powerful in activating the cortex? What happens during the recovery period?

Related experiments in which anatomical pathways were studied during the recovery period indicated that the connectional anatomy of the mature brain is not fixed. Instead, new connections can form, and this plasticity is the basis of the functional recovery. Most of the new connections are those formed by the growth of the terminal branches of the preserved axon arbors into the nearby or adjacent territories of axon arbors that were lost or totally deactivated by the damage to the sensory system. If neurons no longer have a source of activation, active axon branches grow toward these silent neurons until they contact and become capable of activating them. This growth, with the formation of new connections occurs at both subcortical and cortical levels of the sensory systems. Synaptic contacts with nearby axons predominate and occur most rapidly, but more distant connections are possible during long recovery times of months. However, the shorter new connections are likely to be the most useful because they are concerned with similar body parts and thereby have more potential for substituting for lost connections. Longer new connections may lead to misperception rather than compensation. For example, a touch on the face may be also felt on the hand when the deactivated hand cortex is reactivated by inputs from the face. Nevertheless, the experiments show that new connections do form in the parts of the brain that have become unresponsive and nonfunctional, and these new connections are capable of restoring high levels of neural activity. Most importantly, restored levels of neural activity in the cortex and elsewhere can result in the return of lost functions.

Such growth patterns and reactivations are not limited to the somatosensory system. Extensive cortical reactivation and reorganization has been demonstrated after sensory loss in the auditory and visual cortex. Presumably, they occur in all cortical areas, sensory or not, because deactivations following brain lesions are followed by reactivations based on the growth of axon arbors of active neurons so that they replace and compensate for the arbors of inactive neurons. Further evidence for this comes from studies where a portion of the primary motor cortex has been removed.

Consequence of Removal of the Motor Cortex

The primary motor cortex of monkeys is another part of the cortex that has proven useful in studies designed to determine if brain changes accompany behavioral recoveries after damage.

The motor cortex of primates lies in the most posterior portion of the frontal lobe, just in front of areas 3a (for proprioception) and 3b (for touch) of the somatosensory cortex. Although the primary motor cortex receives sensory inputs from somatosensory areas of the cortex, it is not practical to determine the somatotopic organization of the motor cortex by recording from neurons throughout the field to determine their receptive fields. Instead, the motor cortex is “mapped” by using microelectrodes to electrically stimulate (activate) neurons in rows of sites across the field. Typically, short bursts of electrical pulses of the right amplitude will evoke a movement of some part of the contralateral body. The movement may be a gross one, possibly involving movements of fingers, the wrist, elbow, and even the shoulder, but the stimulating current can be lowered until only a small, simple movement remains, such as the twitch of a single finger. When the electrode is moved to another nearby site, and the process repeated, a different small movement is usually obtained. By stimulating sites throughout the primary motor cortex, an overall movement map is obtained. This is possible because a major output of the primary motor cortex is to neurons of the contralateral brainstem and spinal cord where these neurons directly or indirectly activate the muscles that produce movements. The overall movement map resembles the sensory map in the somatosensory cortex (area 3b) in that in both areas the hind limb and foot are represented medially in the motor cortex, the trunk more laterally, the forelimb and hand still more laterally, and the face and tongue most laterally. In the motor cortex, however, we have a map of evoked movements and in area 3b, we have a map of sources of sensory activation. Perhaps as a result, the maps differ in detail. The sensory map in area 3b is systematic such that digits are represented in order with the hand portion of the field, but the motor map does not have this detailed order at a fine grain level of representing movements. There is a hand representation where movements of fingers are evoked, but similar finger movements can be evoked from several, nonadjacent sites, and sites evoking wrist or even elbow movements might be mixed in. Thus, a lesion of part of the hand representation in the motor cortex would remove a mixture of sites related to various hand and forelimb movements, but leave a mixture of such sites in the preserved hand cortex.

As one might expect, even an incomplete lesion of the hand representation in the motor cortex of monkeys impairs skilled hand use. Although crude movements of the hand remain immediately after the lesion, the monkeys have difficulties in picking up small pieces of food or in retrieving them from shallow holes. However, during the weeks to months of post-lesion recovery, the monkeys improve on these tasks, eventually recovering most of their abilities to make skillful finger movements. This recovery is accompanied by a change in the organization of the motor cortex. When electrically stimulated, the remaining hand cortex has changed its organization so that more sites evoke finger movements, compared with wrist or elbow movements. The motor map has been altered in the effectiveness of its output so that more of the remaining cortical hand region is devoted to finger movements. This is at least partly the result of changes in the distribution of the termination of axon arbors of the preserved cortical neurons as they grew into the territories of spinal cord motor neurons that were deprived of many of their activating inputs by the cortical lesion.

These results from lesions of the motor cortex are not likely to be unique to the motor cortex. Lesions of any portion of the cortex will disrupt the functions of the damaged region of cortex and remove outputs to other areas of the cortex and to subcortical structures. These portions of the brain, now deprived of their activating inputs, will not function normally, and impairments in abilities will follow as processing in a network of cortical areas and subcortical nuclei has been disrupted. In addition, neurons in cortical areas and subcortical nuclei that projected to the lesioned cortex will be damaged as they lose their axon terminals that were in the lesion site. Thus, damage to a part of any cortical area is likely to be followed by a sequence of repairs and reorganization, both within partly damaged fields and within fields projecting to the damaged cortex or receiving inputs from the damaged cortex (cortical areas generally both receive inputs and send projections to other cortical areas). Likewise, subcortical thalamic nuclei and other structures with inputs and projections altered by a cortical lesion would reorganize. These reorganizations would be based on the formation of new connections, and reactivated neurons would contribute to behavioral recoveries.

Current research is focused on the use of neurochemical treatments that could enhance the new growth of connections after brain damage or sensory loss, and possibly improve recoveries. In addition, considerable evidence indicates that post-lesion training and experience can improve recoveries by influencing the formation and selection of synapses in damaged systems, thereby promoting the formation of more useful neural circuits. Further research has the potential of revealing those treatments and procedures that are most effective in promoting recoveries.

Jon H. Kaas

See also Cortical Organization; Experience-Dependent Plasticity

Further Readings

Buonomano, D. V., & Merzenich, M. M. (1998). Cortical plasticity: From synapse to maps. *Annual review of neuroscience, 21,* 149–186.

Kaas, J. H. (Ed.). (2001). *The mutable brain.* Amsterdam: Harwood Academic.

Kaas, J. H., Qi, H. X., Burish, M., Gharbawie, O., Onifer, S. M., & Massey, J. M. (2008). Cortical and subcortical plasticity in the brains of rats and primates after damage to sensory afferents in the dorsal columns of the spinal cord. *Experimental Neurology, 209,* 407–416.

CROSS-MODAL TRANSFER

Can a person born blind, upon gaining sight, differentiate between a sphere and a cube using vision alone? In other words, can knowledge previously learned by touch be used to inform vision? This question, first posed by William Molyneux in the late 17th century, concerns cross-modal (across senses) transfer, when information gathered in one sensory modality (e.g., touch) transfers to knowledge in another sensory modality (e.g., vision). Cross-modal (CM) transfer results from stimulus characteristics that are related across the senses. For example, an object with a rough texture has a property (roughness) that is readily picked up by touch, but that can be detected by vision as well, as a visually uneven surface. For this reason, CM transfer is sometimes referred to as intermodal (between senses) equivalence.

Though there may be CM transfers across all senses, most research focuses on transfer between vision and touch, with some research on transfer between vision and audition. This entry describes methods of testing CM transfer, transfer between vision and touch, and transfer between hearing and vision.

Methods

To test CM transfer, adults are asked to recognize by sight (or touch or sound) an object they previously experienced in a different modality. Another method involves training subjects on a task with one modality, then asking them to perform an analogous task in another modality. Performance on the second task is compared with that of subjects who had no training on the first task. Faster learning or better performance of the first group is attributed to CM transfer.

A method typically used with infants involves familiarizing them with a stimulus in one modality until they are habituated to it (no longer show interest), then presenting the same object in another modality paired with a new object. Infants show evidence of cross-modal transfer if they explore the new object longer, as they have already learned about the familiar object in the first modality.

Transfer Between Vision and Touch

Although adults can transfer information from touch to vision and from vision to touch, there are situations in which they do not succeed, or are less accurate. For example, after holding an apple with closed eyes, an adult may be able to identify it when shown a visual array of differently shaped objects. Experimental factors that affect performance include length of exposure, object complexity, and stimulus properties (e.g., length versus shape). Another crucial factor is the degree to which vision and touch can derive the same information about a stimulus, known as perceptual equivalence, which depends on how strongly the stimulus characteristics are related across the senses. For example, if touch and vision were equally accurate at picking up an object's roughness, then roughness would be perceptually equivalent for touch and vision.

Shape

Overall, adults can visually recognize a shape (e.g., a cylinder or a prism) they previously felt and can identify by touch a shape they just saw. However, there is often an asymmetry, such that they are better at CM transfer from touch to vision than from vision to touch. Similarly, newborns can visually recognize a shape they previously felt, at least when they explored it with their right hands, but often fail to transfer the same shape information from vision to touch with either hand throughout the first year of life. The opposite asymmetry is found occasionally during development and depends on how different the test shapes are from one another, as well as the length of time subjects explore the objects.

Texture

Adults and children are generally accurate at transferring information about whether a texture is rough or smooth from vision to touch, as well as from touch to vision. Newborns also show evidence of transfer of texture information between vision and touch. At 1 month of age (the youngest age tested), infants can transfer information about softness and hardness from oral exploration to vision.

Despite the evidence of CM transfer of texture in both directions, when the stimuli tested possess a low degree of perceptual equivalence, an asymmetry emerges at all ages in favor of transfer from vision to touch.

Asymmetry in CM transfer could result because different properties of objects are easily picked up by vision and touch. Because vision is arguably better suited than touch to extract shape information, it might be easier to recognize a previously felt shape by sight than it is to recognize a previously seen shape by touch. Although touch is arguably better at picking up texture information, there is generally not an asymmetry in CM transfer of texture. The different patterns of asymmetry may reflect stronger perceptual equivalence across vision and touch for texture than for shape.

Although CM transfer improves with age, individual differences in CM transfer are apparent from infancy and continue throughout development; infants who display greater CM transfer in infancy also do so later in childhood.

At-risk infants (e.g., premature infants) are worse at CM transfer than are full-term infants, even when they have the same general sensory functioning. At-risk monkeys (e.g., low birth weight) also show poorer performance on CM transfer tasks. Further, monkeys raised in the dark later fail to show CM transfer to the previously deprived modality (i.e., vision), even when sensory limitations are accounted for. Thus, although CM transfer is present at birth before visual experience, its later refinement depends on visual input. Molyneux's blind man would be only partially successful.

Transfer Between Hearing and Vision

Adults are fairly accurate at transferring information between vision and audition, though transfer abilities/asymmetries differ based on the stimulus properties tested. Adults transfer information about intensity between audition and vision in both directions, but for rhythm and complex auditory patterns, they only show evidence of transfer from vision to audition: they are better at discriminating auditory patterns (e.g., Morse code) after training on an analogous visual task but do not transfer auditory training to the visual task. However, with simpler rhythms, 4- to 5-year-old children show transfer in both directions, but 4-month-old infants only show transfer of rhythm from audition to vision. Newborns can transfer intensity from vision to audition (the only direction tested): They respond less to a loud sound after being exposed to a bright light.

Despite the difficulties in making conclusions from diverse lines of research, evidence indicates that some cross-modal transfer exists at birth. Thus, cross-modal transfer may be a fundamental property of our sensory systems, one that contributes to our rich understanding of the world around us.

Ferrinne Spector and Daphne Maurer

See also Haptics; Infant Perception; Multimodal Interactions: Visual–Auditory; Multimodal Interactions: Visual–Haptic

Further Readings

Marks, L. E. (1978). *The unity of the senses: Interrelations among the modalities.* New York: Academic Press.

Picard, D. (2006). Partial perceptual equivalence between vision and touch for texture information. *Acta Psychologica, 121,* 227–248.

Rose, S. A. (1994). From hand to eye: Findings and issues in infant cross-modal transfer. In D. J. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception* (pp. 265–284). Hillsdale, NJ: Lawrence Erlbaum.

Sann, C., & Streri, A. (2007). Perception of object shape and texture in human newborns: Evidence from cross-modal transfer tasks. *Developmental Science, 10,* 399–410.

CULTURAL EFFECTS ON VISUAL PERCEPTION

Mainstream psychology has generally assumed that psychological processes are universal and that the main role of psychology is to investigate these universal aspects of human beings. Visual perception, attention, and even visual illusion have, therefore, been understood mainly through the underlying optical mechanisms and characteristics of visual information hardwired in the human brain and shared by human beings in general.

During the last couple of decades, however, increasing numbers of cross-cultural studies have empirically reexamined this theoretical assumption and advocated an alternative view of human psychology in which culture and human psychological processes are considered to mutually influence one another. This entry reports some recent attempts to reexamine the so-called universal systems of visual perception and discusses the possibility of cultural influences on perception as evidenced by cultural variations in optical illusion, in color perception, in visual attention, and in brain functioning that governs visual attention.

Cultural Effects on Visual Illusion

In the literature of psychology, optical illusion is often used as evidence of human universals in perception. One of the most famous optical illusions is the Müller-Lyer illusion (see Figure 1), in which people perceive that a line segment ending in inward-pointing arrows is longer than a horizontal line segment ending in outward-pointing arrows. This seemingly universal phenomenon, however, has been tested cross-culturally, and the results indicate cultural variations in the magnitude of illusion. For example, Murray Islanders in Melanesia and members of the Toda tribe in India showed significantly smaller errors than do their British counterparts in judging the relative lengths of the lines. Similarly, extensive cross-cultural studies of 17 societies—including a variety of African agricultural and hunter-gatherer cultures, an Australian Aboriginal foraging culture, a tribe of Filipino horticulturalists, and midwesterners in the United States—show that the degree of illusion is much stronger among U.S. residents. Furthermore, children in some cultures (e.g., hunter-gatherers from the Kalahari Desert) were completely immune to the Müller-Lyer illusion. The findings suggest that individuals who grew up in certain visual environments are not vulnerable to the Müller-Lyer illusion.

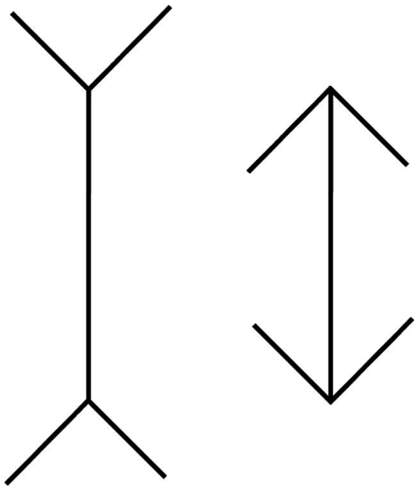


Figure 1 Müller-Lyer Illusion

Source: General Social Survey, National Opinion Research Center (2000).

Various studies have proposed hypotheses to examine the main causes of cultural variations in susceptibility to this illusion. The carpentered environment hypothesis, for example, suggests that people developmentally acquire perceptions of a three-dimensional world in accordance with their experiences with the surrounding environment. In Western industrialized societies, individuals' depth of field is founded on the structure of rooms, houses, and furniture consisting of vertical and horizontal lines with corners in a variety of angles. People in these societies associate acute angles with nearby objects (such as the corner of a rug), and obtuse angles with somewhat more distant views (such as the intersection of two walls and a floor). Once they acquire this specific perceptual pattern in the three-dimensional world, they apply the same rules even when they observe the visual representation in the two-dimensional field. The Western perspective in art is a good example. In Western perspective, objects close to the viewer are drawn larger and are characterized by acute angles, and objects farther from the viewer are drawn smaller and feature obtuse angles. For this reason, Westerners perceive a line ending in inward-facing arrows to be farther away (and therefore actually longer) than it appears. However, in cultures where structures are built using less angular shapes, people have fewer opportunities to interpret the relationships between lines and angles in their perceptual world. The carpentered environment hypothesis thus helps explain why people from some cultures are less susceptible to the Müller-Lyer illusion than Westerners.

Drawings using Western perspective have been used to study cultural variations in depth-of-field perception. For example, one study examined how children and illiterate adult laborers from an African Bantu tribe interpreted the image of a large hunter aiming a spear in the direction of two animals, an elephant and an antelope (see Figure 2). The elephant was closer to the hunter from a non-Western two-dimensional point of view, but because it was drawn smaller than the antelope and the hunter, it would be farther away according to Western perspective. The antelope was farther away from the hunter in the two-dimensional view, but because of its size, it would be considered closer to the hunter in Western perspective. Therefore, although Westerners would be expected to perceive that the hunter was aiming at the large prey, the Bantus perceived the same image without the depth of the field, and for this reason they reported that the hunter was aiming at the smaller prey, which they believed to be closer to the hunter.

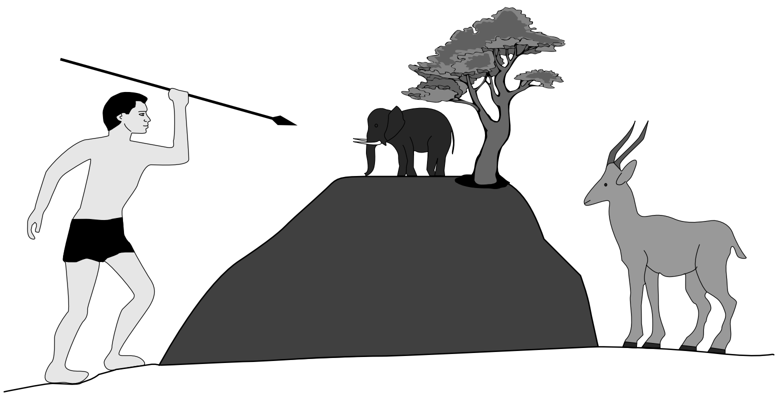


Figure 2 The Image Used in William Hudson's (1960) Experiment

Source: Hudson, W. (1960). Pictorial depth perception in sub-cultural groups in Africa. *Journal of Social Psychology, 52,* 183–208. Reprinted with permission of the Helen Dwight Reid Educational Foundation. Published by Heldref Publication, 1319 Eighteenth St., NW, Washington DC 20036–1802. Copyright © 1960.

The causes of cultural variations in the effects of optical illusions need to be studied further. Current findings, however, suggest that humans' susceptibility to optical illusion may depend heavily on their visual experiences in the environment and on culturally shared interpretations of visual information.

Cultural Effects on Color Perception

Colors accentuate our daily life. But do all people perceive color in exactly the same way? Color is an excellent stimulus for use in scientific investigation because although the spectrum is physically defined, color perception entails psychological processes. One line of research provides evidence that supports the universality of color perception. A cross-cultural study of 98 societies suggests that 11 colors (red, orange, yellow, green, blue, purple, pink, brown, gray, black, and white) are universally recognized, even in the absence of color terms corresponding to these colors. For example, although the Dani people of Irian Jaya on the island of New Guinea had only two color terms (dark and light), they were able to quickly distinguish among the basic colors.

Moreover, researchers who investigate the relationships between colors and color terms advocate an evolutionary account of the development of color terms. That is, the appearance of color terms is predictable according to the number of color terms in a language. If the language has only two color terms, these always turn out to be black (dark) and white (light). However, if the language has three color terms, the third term will correspond to the red end of the spectrum. The fourth and fifth color terms will be green and yellow (or yellow and then green), the sixth color will be blue, and the seventh color will be brown. The last color terms to appear will be purple, pink, orange, and gray, not necessarily in that order.

New findings challenge the idea that color perception is universal among humans, however. For example, the Dani performed relatively poorly in tasks requiring them to remember the basic colors; their retention rate was much lower compared with that of English speakers. Another study investigated color perceptions of Berinmo speakers in East Sepik, Papua New Guinea, whose language contains five basic color terms. The Berinmo's language distinguishes between *nol* (a kind of bluish green) and *wor* (a kind of yellowish green) but not between blue and green. In this study, English and Berinmo speakers were asked to view and remember color chips representing colors spanning either the blue-green boundary or the nol-wor boundary. The Berinmo speakers' performance of the color retention task was better with regard to colors on the nol-wor boundary. The English speakers did better at remembering colors on the blue-green boundary. These findings suggest that linguistic color terms do affect the ability to remember specific colors.

Does language play a role in facilitating or inhibiting our perceptual processes? A study of Russian and English speakers examined whether the existence of a language category influences color perception. Russian has two independent color terms to represent blue: light blue (*goluboy*) and dark blue (*siniy*), whereas English speakers usually distinguish these colors by adding adjectives (light or dark) to the base term *blue*. English and Russian speakers were presented with a variety of blue color chips that were slightly different from each other in hue and saturation, and the task was to discriminate which one corresponded to the target chip. For example, a blue chip was presented as a target stimulus; subsequently, two alternative blue chips were presented, and the participants were asked to select which of the alternatives was identical to the target chip. Russian speakers were quicker to discriminate between two colors when they fell into different categories (goluboy and siniy) than when they were from the same color category. However, these differences did not provide the same advantages for English speakers. This experiment demonstrated that categories in a language affect participants' performance in simple perceptual color tasks. This line of research suggests that culturally shared ways of naming colors may influence speed of color perception.

The issue of universality versus cultural variability in color perception is still controversial. One position holds that language and cultural conventions do not affect color perception; the other maintains that color perception is subject to arbitrary, culturally defined color terms. It has recently been suggested that both positions are partially true. The color terms used in a given culture do influence retention, learning, and ongoing processes of color discrimination. But there is still about 75% overlap in how cultures draw boundaries around color terms. For example, the boundaries of the color terms used by the Berinmo speakers mentioned earlier are similar to those of the five color terms used by Himba speakers in Namibia, although the ecologies and economies of the two groups diverge significantly. Thus, there seem to be universal constraints regarding categorization of colors.

Cultural Effects on Visual Attention

Another line of research has examined whether culture can affect attentional processes. Research on perception and cognition indicates that there are systematic cultural variations in attention between people in East Asian societies (e.g., China, Korea, and Japan) and Western societies (e.g., Canada and the United States). East Asians, who holistically attend to the entire field and relationships between objects, are more context sensitive than Westerners, who analytically focus on salient objects and can easily separate target objects from the context.

A research group used the rod and frame test to compare the levels of context sensitivity of U.S. and Chinese participants. This task used a device consisting of a square frame box of a certain depth, with a rod situated at one end of the box. The participant sat at the other end of the device with his or her chin on a chin-rest, observing the rod through the box. The participants were asked to manipulate the position of the rod until they subjectively thought the rod was perfectly vertical. The experimenter then manipulated the angle of the frame box, and the participant tried again to position the rod vertically. When the angle of the frame was vertical, the frame could serve as a reference point for the position of the rod. When the frame was tilted, however, the judgment of participants who could not ignore the frame would be hindered. The Chinese participants made more errors than the U.S. participants did, suggesting that the Chinese were more sensitive to contextual information and therefore more influenced by the angle of the frame. U.S. participants were able to detach the task from the influence of the angled frame, which suggests that they were relatively immune to the contextual cue.

Are there cultural differences in how attention affects memory for objects in scenes? In an object recognition task, U.S. citizens and Japanese were presented with pictures of wild animals in natural settings. The same participants were then shown pictures of the original animals, as well as new animals, and asked to identify which animals they had seen previously. In this part of the study, the combination of animals and backgrounds was manipulated: Half of the original animals were presented with their original backgrounds, and the rest with completely new backgrounds. Although the task was to identify the animals, the results indicated that, compared with those from the United States, Japanese participants were less able to recognize previously seen animals—especially when they saw them against the novel backgrounds. These results suggest that the Japanese encoded the background information in the images they saw in first part of the study and had more difficulty detaching the target animal from the context.

What are the underlying mechanisms of cultural variation in patterns of attention? And to what extent do sociocultural factors influence our patterns of attention? Recent findings in psychophysiology and neuroscience provide evidence that culture deeply influences attention. Results of an eye-tracking study indicated that East Asians were more likely than were Westerners to allocate their attention to the surrounding information. When given the aforementioned animal recognition task, Chinese participants made more saccadic (rapid nonfocused) eye movements to the background scenes than did those from the United States, even though the task was to evaluate the target objects. These results suggest that context-oriented attention is deeply internalized among East Asians, and for this reason, they cannot help referring to contextual information even when they do not have to. These findings are further supported by a study that measured brain activations during a similar object versus background task. In this study, magnetic resonance imaging revealed that when identifying objects and their locations, more brain regions relating to object information processing were activated in U.S. participants than in Chinese participants. Another cross-cultural study measured activity in the brain area that processes object recognition. There were no identifiable differences between Singaporean college students and their U.S. counterparts; however, Singaporeans 60 years of age and older showed less activity in that brain area than did U.S. residents in the same age group. These results suggest that even the neural circuitry for attending visual scenes is affected by culturally influenced information processing over the long term.

What are the causes of this systematic cultural variation? Researchers in general maintain that people internalize a specific pattern of attention through their experiences of living in a given cultural environment. Some researchers maintain that exposure to culturally biased visual representations such as paintings, drawings, and even photographs facilitate the internalization of a specific pattern of attention. For example, East Asian painting masterpieces were found to be more context rich than are their Western counterparts. Furthermore, when asked to draw scenic images, contemporary members of East Asian cultures were more likely than Westerners were to draw context-rich images. Other studies suggest that East Asian cultures emphasize a sense of interdependence, whereas Western cultures emphasize a sense of independence regarding interpersonal relationships and reasoning styles. Such cultural values may encourage culturally adaptive patterns of attention—for example, making Westerners more likely to see independent objects in the scenes and Easterners more apt to see relationships and contexts that surround the objects. In sum, these findings support the notion that cognitive experiences in the real world influence the processes of our visual systems.

Implications of Research on Culture and Perception

Current research provides evidence of cultural influences on perception. These effects have a variety of implications for social, cultural, and personality psychology, as well as for cognitive psychology. The reasons are threefold. First, the cross-cultural examination of human perception allows us to examine in what ways, and to what extent, our perception is flexibly structured and influenced by systems associated with sociocultural experiences. Some researchers maintain that basic visual processing exists independently of socioculturally shared beliefs. Their findings suggest that the physical and structural systems of visual perception are sufficient for understanding human perception. However, under the rubric of "new look psychology," which emphasizes influences of beliefs and values on visual perception, researchers maintain that our perceptions, even perceptions of so-called neutral stimuli, are fully influenced by our knowledge structures, which in turn are based on our experiences. The underlying processes have not been fully investigated, however, and further research is necessary.

Second, social and cultural psychologists who have identified cultural variation in social cognition—such as causal attribution, self-perception, judgment, inference, and categorization—have long awaited more objective measurements than previously existing quasi-experimental and quasi-survey data collection, which was based mainly on participants' self-reports. Current technological advances allow cross-cultural researchers to scrutinize underlying processes of these variations in human behaviors.

Finally, the theoretical frameworks of perception research do not sufficiently account for the functions of emotions, motivation, and psychological states. Since the emergence of new look psychology, however, substantial numbers of studies have suggested that such factors play an important role in perceptual processes. Again, the findings of cultural influence on perception mutually accelerate further investigation into the complexity of human perception.

Takahiko Masuda

See also Aesthetic Appreciation of Pictures; Attention and Emotion; Color Perception; Eye Movements and Action in Everyday Life; Eye Movements During Cognition and Conversation; Individual Differences in Perception; Nonveridical Perception; Social Perception; Visual Illusions; Visual Scene Perception

Further Readings

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution.* Berkeley: University of California Press.

Chua, H. F., Boland, J., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences USA, 102,* 12629–12633.

Kay, P., & Regier, T. (2007). Color naming universals: The case of Berinmo. *Cognition, 102,* 289–298.

McCauler, R. N., & Henrich, J. (2006). Susceptibility to the Muller-Lyer illusion, theory-neutral observation, and the diachronic penetrability of the visual input system. *Philosophical Psychology, 19,* 1–23.

Nisbett, R. E., & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Sciences USA, 100,* 11163–11175.

Park, D., & Gutchess, A. (2006). The cognitive neuroscience of aging and culture. *Current Directions in Psychological Science, 15,* 105–108.

Roberson, D., Davis, I., & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General, 129,* 369–398.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences USA, 104,* 7780–7785.

Witkin, H. A. (1967). A cognitive-style approach to cross-cultural research. *International Journal of Psychology, 2,* 233–250.

CUTANEOUS PERCEPTION

The skin, far from being just a passive wrapping for the body, provides a wealth of capabilities that combine to allow for extraordinarily complex patterns of perceptual experience. Although cutaneous perception might be taken for granted by most persons, for individuals with visual or auditory disabilities, their impression of the world can depend heavily on their senses of touch. Cutaneous perception results from combinations of responses from skin receptors, evoked by mechanical and thermal stimuli, and, occasionally, chemical and painful events. Historically, there has been some question about the structures and mechanisms that mediate these percepts, partially because it is so difficult to isolate the suspect components, and because some sites (such as the cornea of the eye) are sensitive to touch, temperature, and pain, yet do not possess specialized structures. Nevertheless, converging data from anatomy, physiology, and experiments using "psychological dissection" have led to strong contemporary models of underlying structural-functional relationships.

Mechanical stimuli include static pressure, movement—such as stroking and vibration—and even skin stretch. Thermal stimuli can result from warming or cooling shifts in skin temperature, with extremes that produce pain. Even chemical stimuli, such as pain-relieving salves, generate numerous sensations, including cooling from menthol and heating and irritation from capsaicin. And electricity, the great nonspecific stimulus, can evoke similar perceptual experiences, bypassing receptors to activate nerve fibers directly, mimicking sensations produced by normal ("adequate") stimulation of the skin. To complicate matters further, these qualities can combine to evoke complex illusory percepts—for example, cold pressure stimuli can feel wet (like touching mercury), and the perception of movement can be produced by a rapid sequence of touches.

Cutaneous perception can result from passive contact with static or moving point-like ("punctate") stimuli, such as a mosquito lighting on our arms, or extended 2-dimensional surfaces, such as sandpaper, the tines of a comb, or even dense vibrotactile displays like the Optacon, a machine that blind people use to read print. Similarly, we are sensitive to changing stimuli, such as the warming of a coffee cup, and are aware, through the whole body's surface, when the ambient room temperature drops several degrees. More complex percepts can be evoked from active exploration of simple or multidimensional stimuli, as when we try to identify a Braille character, determine a tomato's ripeness, or assemble a wristwatch. Sensitivities to stimuli vary across the body, leading to different perceived qualities. This situation holds because the structures that subserve tactile experience, the cutaneous receptors, differ in type and density from one site to another. For example, the elbow has recently been empirically shown to be a more sensitive site for thermal stimuli (as mothers, testing their babies' bathwater, have always known) than other areas of the arm.

The skin has two broad divisions—glabrous (smooth) skin, such as the fingertips, and hairy skin, which covers most of the human body. Distinctions between these two skin types include the presence or absence of hairs and the intricate labyrinthine fingerprints. Most research studies of the skin and its capabilities have concentrated on glabrous sites, particularly the fingers and hands. From these experiments, physiological models of cutaneous perception that relate particular characteristics of tactile experience, such as roughness or stretch, have been proposed (Joel Greenspan and Sandy Bolanowski provide a detailed history and description). Because of differences in the receptor populations between the skin types, these models based on glabrous skin should be extrapolated to areas such as the limbs or trunk, only with great caution. This entry discusses cutaneous perception in relation to intensity, space, and time and describes real and virtual tactile surfaces and environments.

Cutaneous Perception and Intensity

Research has shown that tactile perception of simple points, or "asperities," can be extraordinarily acute—we can feel (and localize) 1 micrometer "bumps" on an otherwise smooth surface with our fingertips. Our experience of feeling imperfections along the surface of an automobile or piece of furniture attests to this ability. Place that bump into motion by vibrating it, and our sensitivity can improve a great deal, under certain conditions. Vibrating an area on the fingertip the size of a pencil eraser at a frequency of about 250 hertz (Hz) can be felt at signal amplitudes of much less than a micrometer. The fingertips, sometimes characterized as the "retina" of the skin, are the most sensitive to vibration. Move that stimulator to the palm of the hand, the wrist, forearm, or chest, and sensitivity drops by a factor of as much as 100. One of the underlying mechanisms for this sensitivity gradient is the reduction in the number of receptors and changes in receptor types. For example, Roland Johansson and Åke Vallbo report that there are more than 130 Meissner's corpuscles per square centimeter (cm^2) in the skin of the index fingertip, whereas at the base of the thumb, there are fewer than 30/cm^2, and they don't seem to exist at all in hairy skin.

The maximum intensity that can be felt depends on a number of stimulus conditions, including site, frequency, contactor size, and age. Usually pain or tissue damage defines the upper limit, but a usable dynamic range between just noticing a stimulus and a comfortably "loud" level can be as much as 10,000:1. Despite this large range, if we wanted to use tactile signals—say, in a cardiac emergency code indicating a range of importance from "Check your blood pressure" to "Call 911!"—cognitive limitations restrict the number of useful intensity levels to three or four, even though we can discriminate many more differences when directly compared. Roger Cholewiak consulted on this kind of problem in the development of an implanted cardiac monitor, the AngelMed Guardian, in which subcutaneous tactile feedback is used to warn the user of the severity of an identified condition. Finally, Joseph Stevens and his colleagues have quantified changes over body loci that occur with aging. These are generally attributed to the reduction in number and "health" of the most sensitive touch structures in elderly persons.

Cutaneous Perception and Space

Of the spatial modalities, touch falls between vision and audition in its acuity, being less precise than vision but more precise than hearing. It is not difficult to locate an insect on the arm because it bends hairs while it walks about. Generally, the ability to localize vibrations on the 2 square meters (m^2) of the skin can be quite good, as long as they can be felt. This ability has been tested empirically with both active and passive presentations of stimuli. Active exploration mimics the typical way we use our skin in everyday life ("haptics"). Passive stimulus presentations, however, allow the researcher to control the signal more precisely, but at the expense of losing the richness of kinesthetic and motor feedback that enhances "everyday" spatial percepts. In the same way that sensitivity to a stimulus varies over the body, so does our ability to localize an event: Touch the fingertip lightly with a pencil point and it will be felt every time; on the back of the hand it will be felt often, but not always, whereas on the chest a light touch might be missed at many loci. A sidelight of this demonstration is to attend to the "coolness" of the tip. On the back of the hand most touched points will be felt as neutral, but occasionally, "cold spots" will brightly announce their presence. These demonstrate that the distribution of cutaneous receptors is neither dense nor uniform. There is an interaction between this punctate sensitivity and perceived intensity: The skin's sensitivity to warmth (as well as to higher-frequency vibrations) depends on the area of stimulation. Specifically, the larger the region warmed or the size of the contactor, the "louder" the sensation, a characteristic called *spatial summation.*

The ability to distinguish whether one or two points have been touched depends on how far apart they are, increasing from about a millimeter on the fingertip to several centimeters on less-sensitive areas such as the abdomen or thigh. Interestingly, there are certain "anchor points" near which localization is better. Although the limb joints serve this function, the midline of the body—front and back of the trunk—have recently been shown by Roger Cholewiak and his colleagues to anchor near-precise localizations. Finally, as Stevens and his colleagues have shown, spatial acuity deteriorates with age, as do many perceptual functions.

A one-dimensional stimulus such as a vibrating point might be employed for a kind of tactile Morse code, or to signal the presence of an event, but the temporal characteristics of the skin limit transmission rates for complex streams. More useful information can be communicated to a person with two-dimensional displays, such as Braille cells, incorporating spatial information. Experienced Braille readers can read at 60 words per minute (wpm), although 300 wpm rates have been reported. (Visual rates range from 250 to 400–600 wpm.) What limits the processing of tactile patterns? One important factor is *masking,* in which stimuli preceding or following a pattern degrade its processing. This degradation can take the form of changes in sensitivity, or in the ability to recognize the pattern. In the latter case, depending on the relative shapes and the timing between patterns, features can be dropped, added, or distorted. For example, a "P" might be perceived as an "F," or an "H" felt as an "A," as James Craig's extensive work has shown. These interactions typically occur when presentations occur within 200 milliseconds (ms) of one another, regardless of whether they are static or scanned across the finger. Another type of spatial interaction, reviewed by Lynette Jones and Susan Lederman, is related to patterns "drawn" on the skin's surface (*graphesthesia*), and the position of the body part in space. Here, identification of similarly shaped letters (such as b, d, p, q) drawn on the hand, arm, thigh, or forehead can depend on the limb's orientation and "point of view" (egocentric vs. allocentric) taken by the viewer. These data suggest that mobile body sites should be used for tactile displays only with caution. The torso has been chosen to present tactile information about the environment for navigation in cases of sensory disability, or for displays for situation awareness (such as Angus Rupert's aircraft Tactile Situation Awareness System) to augment overloaded "major" senses.

Cutaneous Perception and Time

Regarding temporal acuity, the skin again takes the middle ground, this time being more acute than vision but less acute than audition. Tests of temporal order indicate that there is some parity among these modalities (the chemical senses usually being considered far slower), so that regardless of the stimulus, a separation of about 20 ms is required to identify the order of two events (brief clicks, flashes of light, or taps on the skin). We are also able to detect gaps in prolonged single-frequency vibration or vibrotactile "noise" (where many frequencies are combined), but again, depending on a number of factors, such as age and stimulus intensity, gaps shorter than 250 ms are difficult for observers to appreciate.

Like vision and audition, tactile perception is limited to a narrow range of temporal variation (frequencies). Whereas the other mechanical sense, audition, has a useful frequency range from 20 Hz to about 20 kilohertz (KHz), that of the skin is more limited, from about 20 to 300 Hz. There are instances of low frequency sensitivity, say to swaying of a tall building, but those experiences are often ephemeral and the sensations confused with internal body functions. And, like vision and audition, a tactile stimulus has to stay on for some minimal time before the richness of its qualities can emerge. A pressure pulse (a "touch") can be felt if it is as brief as 2 ms, and increases in perceived intensity with duration, a phenomenon described as *temporal summation.* But not only do stimuli briefer than about 200 ms have to be presented at higher intensities even to be felt, for durations far below than this, vibration will not feel periodic (nor will sound have tonal quality—the "atonal" interval). However, because the more-sensitive skin receptors tend to respond best to transients, even more durative stimuli (either pressure or vibration) won't necessarily be appreciated as being proportionally stronger and might lead to the sensory phenomenon known as adaptation. That is, like the constant pressure of the clothes on our body, prolonged vibration leads to a reduction in apparent intensity.

So, the changes in the several perceptual qualities associated with temporal summation lead to a recommended upper limit for vibrotactile bursts of about 200 ms, beyond which sensation magnitude can fall. There is a similar range of thermal sensitivity (our "physiological zero") that occurs over a limited span of ambient temperatures where we may feel neither warm nor cool, given enough time to adapt. All our sensory systems are tuned to respond to changes, considering constant stimuli less informative: Sitting still provides little information about our clothes, but move the arm and we can become aware of the fabric around our sleeve, if we pay attention. Given these limitations, a vibrotactile Morse code could result in communication with relatively slow transmission rates (Hong Tan and Nat Durlach showed that at most, about 20 wpm can be achieved tactually, the amateur level for acoustic Morse).

Interestingly, when trying to determine whether one or two points have touched the skin, introducing a difference in time can make the task trivial. Even when identifying the orientation of two-dimensional gratings (similar to the tines on a comb), if the fingertip can stroke the surface rather than have it passively touched, the array becomes a spatiotemporal display and the groove orientation becomes immediately obvious. Research has shown that grid orientation tasks clearly show the influences of aging on the ability to distinguish texture, as long as stimuli are passively presented. However, allow the finger to stroke the surface and there is no difference between a 10-year-old and an octogenarian.

Because movement is a change in location over time (spatiotemporal), it has a number of perceptible qualities in those domains, such as direction, distance, and velocity. Greg Essick has shown that movement can be generated on the skin in a variety of ways (such as a brush dragged across the skin or a series of taps on individual vibrators), and that we are good at identifying its direction, unless the movement is very fast. Our perception of extent and "straightness" also depend on velocity: If the sequence is too fast, perceived extent may be foreshortened, but if it is too slow, the path may wander. With appropriate controls, illusory motion (akin to vision's "Phi") can be observed with only two tactile stimuli, such as vibrations at locations separated by 10 cm and 100 ms. A different illusory experience, described by Frank Geldard, is evoked by a sequence of, say, five taps at one site, followed by a sixth at a second site about 10 centimeters (cm) away, with inter-tap intervals of about 50 ms. In this case (sensory saltation), the series will be felt evenly distributed between the two sites. In all of these, the sensations of movement can generate tactile "vectors" within virtual environments for communication systems and appear to have correlates in the central nervous system.

Tactile Surfaces and Environments: Real and Virtual

Because of the interest in applying tactile displays to enhance virtual environments, as well as for sensory substitution and augmentation, it has become important to study the ability of the skin to appreciate physical dimensions of real-world surfaces and structures. As children, we would lay paper on the ground and create patterns with crayons, the tip rising and falling with the underlying surface. This texture was transmitted to cutaneous receptors via the crayon's vibration, and the surface roughness perceived through two primary sources. The vibrotactile information, spatial variation over time, conveys something about a surface's features, but alone does not provide enough information to form the percept of a texture (e.g., a vibrating cell phone does not feel like sandpaper). Adding proprioceptive feedback—information from muscles and joints—gives egocentric knowledge of the relative locations of each body part. The combined information from vibrotactile and proprioceptive sources underlies tactile perception of texture. Additional information, such as from vision, can form an even stronger percept.

Combinations of these spatial and temporal qualities in tactile exploration can make us aware of a number of physical surface qualities, including stiffness, force, and friction. To assess the ripeness of a pear, it's often helpful to squeeze the fruit and feel its stiffness. Stiffness—the force exerted by a surface proportional to the distance it is compressed or stretched—can be perceived because of the mechanical stimuli (static pressure and movement) and the proprioceptive feedback of the joints. Force is a bit more difficult to distinguish from other characteristics because the static pressure sensed by the cutaneous receptors is supplemented by skin stretch and displacement. Some attributes of a surface, such as friction, might be appreciated using vision, but the skin often provides information that would be otherwise imperceptible, such as the stickiness of flypaper. Although vision helps to guide the extremities, haptic information ultimately provides information about the complex forces (e.g., weight and friction) and compliancy of surfaces, as Steven Cholewiak, Hong Tan, and David Ebert have shown. These qualities, as well as texture, are important for appreciating surfaces in everyday and virtual realities.

There is particular interest in using cutaneous perception and haptics to enhance skill learning in many virtual environments such as telesurgery. Telesurgery, performed using a human-controlled robot, relies on surgeons' abilities to map their movements with controllers to the three-dimensional locations of the robot's appendages. It requires long and tedious practice proficiency. Any features that could make the surgery more "life-like" can potentially smooth the transition. By providing haptic force and cutaneous feedback (e.g., texture, pressure, temperature, stiffness), learning time may be reduced and the procedure taught to a wider array of individuals, who may have avoided the technology because of its perceptual sterility.

Roger W. Cholewiak and Steven A. Cholewiak

See also Cutaneous Perception: Physiology; Haptics; Texture Perception: Tactile; Vibratory Perception; Virtual Reality: Touch/Haptics

Further Readings

Angel Medical Systems. (2008). The AngelMed Guardian System. Retrieved August 24, 2008, Angel Medical Systems Product page, http://www.angel-med.com/product.html

Blankenburg, F., Ruff, C., Deichmann, R., Rees, G., & Driver, J. (2006). The cutaneous rabbit illusion affects human primary sensory cortex somatotopically. *Public Library of Science, Biology, 4,* 369.

Cholewiak, R. W., Brill, J. C., & Schwab, A. (2004). Vibrotactile localization on the abdomen: Effects of place and space. *Perception & Psychophysics, 66*(6), 970–987.

Cholewiak, S. A., Tan, H. Z., & Ebert, D. S. (2008). Haptic identification of stiffness and force magnitude. *Proceedings of the Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems,* 87–91.

Craig, J. (2002). Identification of scanned and static tactile patterns. *Perception & Psychophysics, 64*(1), 107–120.

Craig, J., & Sherrick, C. (1982). Dynamic tactile displays. In W. Schiff & E. Foulke (Eds.), *Tactual perception: A sourcebook* (pp. 209–233). Cambridge, UK: Cambridge University Press.

Essick, G. (1998). Factors affecting direction discrimination of moving tactile stimuli. In J. Morley (Ed.), *Neural aspects in tactile sensation* (Vol. 127, pp. 1–54). New York: Elsevier.

Geldard, F. (1982). Saltation in somesthesis. *Psychological Bulletin, 92,* 136–175.

Greenspan, J., & Bolanowski, S. (1996). The psychophysics of tactile perception and its peripheral physiological basis. In L. Kruger (Ed.), *Pain and touch* (pp. 25–104). San Diego, CA: Academic.

Heller, M., & Schiff, W. (1991). *The psychology of touch.* Hillsdale, NJ: Lawrence Erlbaum.

Johansson, R., & Vallbo, Å. B. (1979). Tactile sensibility in the human hand: Relative and absolute densities of four types of mechanoreceptive units in glabrous skin. *Journal of Physiology, 286,* 283–300.

Jones, L., & Lederman, S. (2006). *Human hand function.* Oxford, UK: Oxford University Press.

Rupert, A. H. (2000). Tactile situation awareness system: Proprioceptive prostheses for sensory deficiencies. *Aviation Space and Environmental Medicine, 71*(9 Suppl), A92–99.

Stevens, J., & Choo, K. (1996). Spatial acuity of the body surface over the life span. *Somatosensory and Motor Research, 13*(2), 153–166.

Summers, I. (1992). *Tactile aids for the hearing impaired.* London: Whurr.

CUTANEOUS PERCEPTION: PHYSIOLOGY

A major challenge in neurobiology is to understand how the brain constructs mental images of the world around us. The mental images that arise from the sense of touch are based on continuously changing patterns of electrical activity called action potentials that are evoked in the nerves that innervate the skin, muscles, and joints. The dynamic patterns of action potentials that come from the skin are the basis of cutaneous perception. These patterns are sent to the central nervous system via two main nuclei located in the brain stem and thalamus. Once the information reaches the cortex, it is systematically transformed through several processing stages into an alternate transformed pattern that is matched against previously stored patterns to evoke mental images of objects and surfaces in contact with the skin. The challenge facing neurobiologists is to understand the anatomical pathways and neural circuits that transform the patterns from the initial pattern into the representation that underlies memory, in other words, the challenge is to understand the neural code(s) that underlie behavior.

When exploring and manipulating an object with our hands, we readily appreciate many qualities or features of the object. These features include characteristics such as its size and shape, the texture of the surface, its weight, and dynamic properties, such as whether it is stationary or is moving in our hand. Many studies have shown that our ability to discriminate and identify objects is based on a rapid pattern recognition mechanism. For example, common everyday objects are recognized (typically in less than 3 seconds) without visual input at accuracy rates greater than 96%. In those experiments, subjects typically report that they identified the object using two to three features, such as its size and texture. In addition to be being highly accurate and rapid, the cutaneous system is also extremely sensitive with young adults being capable of detecting vibrations with amplitudes as low as 100 angstroms.

Discovering the neural code(s) that underlie cutaneous perception has been difficult for a number of reasons. First, the sense of touch is composed of multiple sub-modalities with individual features being coded by different afferent types. Somehow, these features are processed and integrated in the cortex to produce a single unified percept of objects. Second, in contrast to vision and audition, the cutaneous receptors are imbedded in a deformable sensory sheet (i.e., the hand) that dynamically changes as the hand moves around in space. Thus, the pattern matching is a dynamic process. Third, sensory inputs play a dual role and are important for both sensation and action. The sensory inputs provide information about the world, and the inputs related to action play important roles in motor control. Fourth, cutaneous perception is tightly linked to higher cognitive functions such as selective attention and drive what aspects of the inputs are given privileged access to perception.

To illustrate the role and complexity of the sensory processing from the hand, imagine what happens during the simple act of grasping and tossing a tennis ball. During the grasping phase, the fingers must spread apart to surround the ball and the hand moves to surround and enclose the ball. After contact, sensory afferents send their outputs to the cortex where they are transformed and matched against stored memories of tennis balls. The sensory inputs also play an important role during manipulation and ensure that the proper grip force is used to prevent the ball from being dropped. Finally, to toss the ball, the motor system must be activated in a way that is highly coordinated with the sensory inputs to allow the ball to be released at just the right time. How this simple task is performed in the nervous system is not known.

This entry first discusses how information related to cutaneous perception is initially encoded in the peripheral afferent nerves that innervate the hand to produce a primary neural representation of the cutaneous inputs. Then the entry discusses the anatomical pathways that carry the information to the cortex and how and where cutaneous information is processed in the cortex.

Primary Afferent Representation

Perception from the hand begins with 13 different types of afferent fibers that innervate the hand. Four of these afferent types respond to mechanical stimuli and provide information about cutaneous inputs. The other 9 afferents provide information about pain (2), temperature (2), itch (1), and hand position (4) and will not be discussed further. Two of the mechanoreceptive afferent types lie close to epidermis and densely innervate the skin, and two afferent types lie deep in the dermis and have much lower innervation densities, which are defined as the number of afferent fibers per square centimeter of skin. One mechanoreceptor is called the slowly adapting type 1 (SA1) afferents. Individual SA1 afferents end in a neurite complex called the Merkel cells, which are specialized epidermal cells located at the tips of the epidermal ridges and enfold the branches of the axon terminal of the SA1 afferents (see color insert, Figure 7). There are about 100 SA1 afferents per square centimeter (cm^2) on the distal finger pads, and the innervation density declines rapidly as you move proximally along the arm toward the body. Sensory neurons and their afferent fibers are characterized by which stimuli cause changes in their firing rates. For the cutaneous receptors, one characteristic is a property called the receptive field (RF), which is defined as the area of skin that causes a change in the firing rate of the afferent fiber. The region on the skin where stimuli cause increases (or decreases) in firing rate are called excitatory (or inhibitory) regions. The RF sizes of SA1 afferent fibers are typically about 2 to 3 millimeters (mm) in diameter and contain one or more central "hot" spots or zones within their RF of high sensitivity. SA1 afferents, like all of the cutaneous afferent fibers, have purely excitatory RFs. The acuity threshold of SA1 afferents, which is defined as the minimum distance that these afferents are able to discriminate one point indented into the skin from two points, is about 1.0 mm. This threshold is nearly identical to the perceptual threshold of human spatial acuity on the finger pad, which demonstrates that these afferents are responsible for two-point discrimination. The SA1 afferents adapt slowly (which is why they are called slowly adapting) to a constant indentation with the baseline firing rate increasing with increasing intensity of the stimulus (probe indented deeper into the skin).

The SA1 system, which consists of the SA1 afferents and their central projections, is the spatial system in touch. When a two-dimensional (2-D) pattern such as an embossed letter of the alphabet is indented into the skin, activity is evoked in the afferents that have part of their RF below the letter while other SA1 afferents remain silent. Because the SA1 afferents densely innervate the skin, the neural activity evoked across the population of afferent fibers is in the same spatial form as the pattern (Figure 1). That is, an isomorphic (i.e., in the same form) neural representation of the 2-D pattern with each afferent fiber acts like a single pixel in the neural image. This peripheral neural image is sent to the brain by the SA1 afferents.

Practically everything that we know about the functional roles of the afferent fibers comes from studies that combine psychophysical measurements in humans with neurophysiological recordings from nonhuman primates. A series of studies showed that the SA1 afferents are not only responsible for spatial acuity and 2-D form but also are responsible for coding features related to texture perception. Those studies show that the perception of roughness (and smoothness) is based on the spatial variation in firing rates among the population of SA1 afferents. Thus, surfaces that cause every other afferent fiber to fire vigorously show high spatial variation in firing and are perceived as feeling rough, whereas surfaces that evoke similar firing rates across the SA1 afferent population are perceived as feeling smooth. Thus, the SA1 system is the spatial system and plays the analogous role in cutaneous perception that the parvocellular system plays in vision.

The second mechanoreceptive afferent type is called the rapidly adapting (RA) afferents. These afferents also branch as they approach the skin with their individual branches embedded in anatomical structures called Meissner's corpuscles (see color insert, Figure 7). Meissner's corpuscles are composed of stacks of laminar disks that are tethered between the sweat ducts and primary ridges of the dermal papillae. Like the SA1 afferents, RA afferents innervate the skin densely (about 150 afferents/square centimeter [cm^2]). However, there are significant differences between these afferent types. RA afferents respond only to the onset and offset of a probe indented into the skin and hence adapt rapidly to the stimulus. These afferents do not respond during the period when the probe is statically indented into the skin. The RF sizes of RA afferents are larger than SA1 afferents (3–6 mm in diameter rather than 2–3 mm) and show uniform sensitivity to an indented probe placed inside the RF. RA afferents show poor spatial acuity and are unable to resolve gaps in gratings spaced more than 3 mm apart.

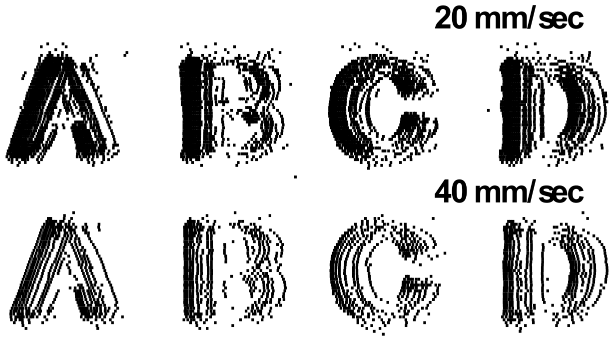


Figure 1 Spatial Event Plot for a Typical SA1 Afferent to Embossed Letters (ABCD) Scanned Across the RF at Two Different Velocities

Note: This plot represents the spatial pattern of input that is sent to the cortex from the population of SA1 afferent fibers.

The RA afferents are highly sensitive to dynamic events such as moving stimuli across the skin and to low-frequency vibrations (below about 60 hertz [Hz]). RA afferents play an important role in grasping by signaling when objects begin to slip in our hands. The working hypothesis is that the RA system, which includes the RA afferent fibers and all of their central projections, is the tactile motion system and is analogous to the magnocellular system in vision.

The third mechanoreceptive afferent type is the Pacinian afferent (PC). This class of afferents is also rapidly adapting, but in contrast to the other SA1 and RA afferents, these afferents do not branch as they approach the skin; instead, they end in the deeper layers of the dermis in a large anatomical structure consisting of concentric layers of fluid filled sacs known as the Pacinian corpuscle (color insert, Figure 7). The multiple layers of the Pacinian corpuscle act as a mechanical filter for low frequency vibrations and cause these afferent types to be extremely sensitive to high frequency vibrations (100–300 Hz). Studies show that PCs can detect indentations as small as 1 nanometer (nm) applied directly to the corpuscle or 10 nm when applied to the skin surface. Because of their extreme sensitivity, the receptive field boundaries for PC afferents are nearly impossible to define.

PCs play the dominant role in coding temporal information that is transmitted from the working end of tools to the hand during tool use. PCs are the temporal system in touch.

The fourth afferent type is the slowly adapting type II (SA2) afferents. These afferents have not been found in nonhuman primates and rodents, and as such, fewer studies have been devoted to understanding their functional role in perception. They are located deep in the dermis. Historically, it was thought that the receptor ending for these fibers was a structure called the Ruffini's corpuscle (color insert, Figure 7); however, recent studies have questioned whether the Ruffini's corpuscles exist in the hand. Thus, the precise receptor ending for the SA2 afferents is at present not known. Neurophysiological studies in humans show that these afferents have large RF sizes spanning about 1 centimeter (cm) of skin and respond with a slowly adapting response to stretching of the skin in tuned directions. The working hypothesis for these afferents is that as the hand is moved to different positions, the pattern of skin stretch on the back of the hand changes in a systematic way. These stretch patterns are encoded in the SA2 afferent population. As such, the population response of the SA2 afferents is thought to code for hand conformation.

Ascending Pathways

The outputs of the peripheral afferent fibers related to cutaneous perception ascend in a fiber tract called the dorsal-column medial lemniscal pathway and synapse on neurons in the dorsal-column nuclei (DCN), which is located in the brainstem (Figure 2). Outputs from the DCN neurons send their projections to neurons located in the contralateral hemisphere in a group of neurons called the ventroposterior complex (VPC) of the thalamus. Little is known about how the neurons in these intermediate stages process information in the primate. However, neurons in the ventroposterior lateral nucleus (VPL), which are the neurons in the VPC that receive cutaneous inputs from the hand, have small receptive fields that are similar in diameter to the RFs of the peripheral SA1 and RA afferents. This suggests that information is minimally processed at these subcortical stages.

Cortical Processing of Tactile Information

From the thalamus, neurons in VPL project in two directions (Figure 2). The main dense projection is to the primary somatosensory cortex (SI), which is composed of four separate areas that are given the names 3a, 3b, 1, and 2. The other projection, which is much smaller, projects directly to the second somatosensory cortex (SII). The small projection to the SII cortex is not well understood, but one hypothesis is that it is the pathway responsible for processing vibratory information because many of the neurons in this pathway have PC-like responses.

Area 3a is located deep in the central sulcus (CS). It receives its inputs from a particular part of VPL that contains neurons that respond best to arm position and movement. Consequently, area 3a is thought to be involved in coding information about body position and movement, which is also called proprioception or the sixth sense. Little is known about how information is coded in area 3a; however, it is becoming a region of interest as a possible cortical location where stimulating electrodes could be placed in humans with prosthetic arms to provide artificial feedback of proprioceptive input.

Area 3b has been studied extensively. It is located in the posterior bank of the central sulcus, and neurons in area 3b respond exclusively to cutaneous inputs. The RFs of neurons in area 3b are large, confined to a single finger pad, and highly elaborate. In contrast to the RFs of the peripheral afferents, the RF of 3b neurons are composed of an excitatory region surrounded by one or more regions of inhibition. Stimuli placed in the excitatory part of the RF cause the neuron to increase its firing rate, whereas stimuli placed in the inhibitory part cause a decrease in firing rate. The receptive field structures that are observed in 3b are similar to the RFs that have been described of simple cells in the primary visual cortex. As in the visual system, these RFs are feature selective and selectivity responsive to spatial features such as the orientation of bars indented into the skin. The minimum angle that can be discriminated by the population response of area 3b neurons is about 20 degrees, which is identical to psychophysical thresholds of orientation discrimination of humans—suggesting that

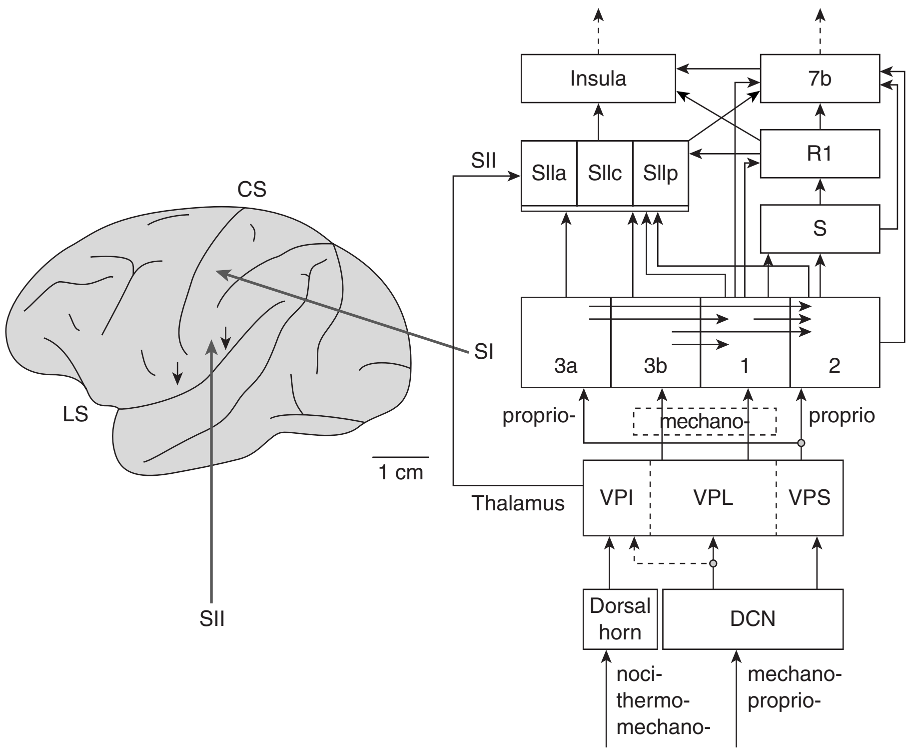


Figure 2 Block Diagram Showing the Flow of Information Through the Somatosensory System

Notes: Outputs from the peripheral afferents related to cutaneous inputs first synapse on neurons in the dorsal column nuclei (DCN). These neurons send their main projections to the ventroposterior lateral nucleus of the thalamus (VPL), which in turn projects to the four areas of the SI cortex and has a minor projection to the SII cortex. The four areas of SI then project to SII and to areas 5, retroinsular (Ri), and 7b. SI is located in the postcentral gyrus of the central sulcus (CS). SII is located in the upper bank of the lateral sulcus (LS). VPI and VPS stand for ventroposterior inferior and superior.

orientation discrimination occurs in area 3b. Discriminating features such as orientation is the first initial step in the transformation of spatial information into the central representation that underlies form processing. These findings suggest that mechanisms underlying form processing in vision and touch are similar and that area 3b is part of the SA1 system for coding spatial form and texture.

Area 1 receives both direct inputs from the thalamus and indirect inputs from area 3b (Figure 2). Like area 3b, area 1 neurons have RFs consisting of excitatory and inhibitory subregions, and area 1 neurons also show orientation-tuned responses. However, several lines of study suggest that it lies further along the processing pathway for form and motion. First, the RFs of area 1 neurons are larger, more complex than 3b neurons, and typically span multiple fingers. This is important because it suggests that area 1 performs further computations on the inputs from area 3b. Furthermore, many area 1 neurons are highly sensitive to stimulus motion. A working hypothesis is that area 1 neurons code for shapes moving across multiple digits.

The RF of area 2 neurons are different from the other parts of SI. They are much larger and typically respond to both proprioceptive and mechanoreceptive inputs. Animals that have area 2 ablated are unable to discriminate large shapes, and neurophysiological and imaging studies suggest that neurons in area 2 respond selectively to objects that differ in three-dimensional (3-D) shape. The working hypothesis for this area is that it is where the initial representations of features related to the 3-D size and shape of objects is generated and plays an important role in processing sensory input for action.

Beyond Primary Somatosensory Cortex

The four areas of the SI cortex project in two directions. One projection is to areas 5, Ri, and 7b (Figure 2). This projection is the action pathway and is important for processing sensory inputs for coding where the body is in immediate personnel space, to directing attention, and to guiding the hand to targets in space.

The other projection is from SI to SII (Figure 2). Animals that have SII ablated are unable to recognize the shape of objects, which means that this projection lies along the pathway underlying tactile object recognition. Many studies now support the idea that SII is composed of multiple areas—SIIa, SIIc, and SIIp in nonhuman primates (Figure 2), and parietal operculum 1 (OP1), OP2, OP3, and OP4 in humans. The different areas in SII respond differently to tactile stimuli. Specifically, the anterior and posterior parts of SII (SIIa and SIIp) respond well to both cutaneous and proprioceptive inputs (such as neurons in area 2). These areas flank a central area of SII (SIIc) that contains neurons that respond mainly to cutaneous inputs. SII neurons have much larger and more elaborate RF than SI neurons do, and many of the neurons (about 30%) show selective responses to oriented bars indented in the skin. SII neurons typically have receptive fields that span multiple pads on the fingers with some pads having orientation-tuned responses and others pads being driven by the stimulus but not being tuned. The diverse set of RF structures that are observed in the SII cortex suggests that this area of the cortex is most likely coding features of large objects that span multiple fingers and is important for processing and integrating information related to the representation of 3-D shapes.

Processing of Touch Beyond SII Cortex

The processing of tactile information beyond SII is poorly understood; however, other areas of the cortex have been shown to be important for processing information about form and vibration. Of particular importance are studies by Ranulfo Romo and his colleagues who have systematically mapped the neural pathways that underlie the decision process in animals trained to discriminate whether vibrations of different frequency presented in sequence were the same or different. Further, neural imaging and studies using transcranial magnetic stimulation, or TMS, in humans suggest that other areas of the cortex play important roles in processing tactile form and texture. Candidate areas include the anterior parts of the intraparietal cortex (IPA), supramarginal gyrus, right intraparietal sulcus (pIPS), and possibly visual areas. How the information is processed in these areas is not known.

The mechanisms of form processing are not understood, but a growing body of evidence suggests that it involves a number of processing stages. The process begins with the activation of the four cutaneous afferents that innervate the glabrous skin of the hand. In the periphery, each of the four types of afferents is responsible for encoding different local features, but in the cortex, the segregation of function appears to be distributed across different hierarchal processing stages with the higher stages responding to more global aspects of the inputs and being influenced by the cognitive state of the animal. About 90% of the neurons in SII are affected by the animal's focus of attention. Beyond SII, the tactile representations are merged with inputs from the other senses to produce a central representation that can be matched to stored memories. Where this matching process occurs is not known.

Steven Hsiao

Supported by NIH grants NS34086 and NS18787.

See also Cutaneous Perception; Pain: Physiological Mechanisms; Proprioception; Vibratory Perception

Further Readings

Bensmaia, S. J., Denchev, P. V., Dammann, J. F., III, Craig, J. C., & Hsiao, S. S. (2008). The representation of stimulus orientation in the early stages of somatosensory processing. *Journal of Neuroscience, 28,* 776–786.

Edin, B. B., & Abbs, J. H. (1991). Finger movement responses of cutaneous mechanoreceptors in the dorsal skin of the human hand. *Journal of Neurophysiology, 65,* 657–670.

Fitzgerald, P. J., Lane, J. W., Thakur, P. H., & Hsiao, S. S. (2006). Receptive field properties of the macaque second somatosensory cortex: Representation of orientation on different finger pads. *Journal of Neuroscience, 26,* 6473–6484.

Gardner, E. P., Babu, K. S., Reitzen, S. D., Ghosh, S., Brown, A. S., Chen, J., et al. (2007). Neurophysiology of prehension. I. Posterior parietal cortex and object-oriented hand behaviors. *Journal of Neurophysiology, 97,* 387–406.

Hsiao, S. S. (1998). Similarities between touch and vision. In J. W. Morley (Ed.), *Neural aspects of tactile sensation* (pp. 131–165). Amsterdam: Elsevier.

Hsiao, S. S., Johnson, K. O., & Yoshioka, T. (2003). Processing of tactile information in the primate brain. In M. Gallagher & R. J. Nelson (Eds.), *Comprehensive handbook of psychology: Vol. 3. Biological psychology* (pp. 211–236). New York: Wiley.

Johnson, K. O., Hsiao, S. S., & Yoshioka, T. (2002). Neural coding and the basic law of psychophysics. *Neuroscientist, 8,* 111–121.

Romo, R., & Salinas, E. (2003). Flutter discrimination: Neural codes, perception, memory and decision making. *Nature reviews. Neuroscience, 4,* 203–218.

Sripati, A. P., Yoshioka, T., Denchev, P., Hsiao, S. S., & Johnson, K. O. (2006). Spatiotemporal receptive fields of peripheral afferents and cortical area 3b and 1 neurons in the primate somatosensory system. *Journal of Neuroscience, 26,* 2101–2114.

Steinmetz, P. N., Roy, A., Fitzgerald, P. J., Hsiao, S. S. , Johnson, K. O., & Niebur, E. (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature, 404,* 187–190.

D

DECISION MAKING, PERCEPTUAL

One can define *perceptual decision making* (PDM) as a categorical choice based on sensory information. For example, did you just hear an airplane flying by or was it a helicopter? PDM can be contrasted with economic, emotional, reward-based, or value-based decision making, in which the decision is driven by the desirability of an option. For example, you see two identical-looking boxes in front of you. From past experience, you know that you are less likely to find candy when opening the right box. You therefore choose the left box. Because perceptual decisions can only be studied when the choice is reported by the subject, studies of PDM usually also involve a verbal report or triggering an action (pushing a button, for example).

Perceptual decisions can be classified based on their complexity. The simplest type of perceptual decision is deciding whether a stimulus has been present or not (stimulus detection). "Did you perceive a flash of light?" would be an example. More complex perceptual decisions involve making a choice between two categories (binary choice). Examples would be "Did you see a cat or a dog?," "Was the second tone louder or softer than the first one?," or "Did you feel a touch on your left or on your right arm?" Even more complex perceptual decisions involve choices between multiple categories (for example, "In which of the four cardinal directions was the object moving?").

The goal of studying PDM is to understand how humans and animals process sensory information for categorizing stimuli and triggering appropriate actions. This involves understanding the choice behavior itself as well as the neural mechanisms leading to this choice behavior. Traditionally, PDM has been studied primarily by psychophysicists and (mathematical) psychologists, but more recently, it has also become a popular area of research in neuroscience. Interdisciplinary approaches that combine behavioral and physiological experimental techniques with mathematical modeling are aimed at understanding the mathematical algorithms underlying PDM (the "recipes" for how to come up with an appropriate choice) as well as their neural implementation (how these computations are performed by the brain).

This entry first summarizes behavioral aspects of PDM, continues with a discussion of neural correlates, and finishes with a review of theoretical concepts.

Behavioral Findings

Accuracy Effects

For a stimulus to be perceived, its intensity needs to exceed a critical level. Thus, it is usually observed that stimuli with intensities well below some critical level are never detected, stimuli with intensities well above this critical level are always detected, and stimuli with intensities close to this critical level are sometimes detected and sometimes missed. The probability of detecting a stimulus as a function of the stimulus intensity is referred to as the *psychometric function.* The stimulus intensity at which the stimulus is detected with a probability of 0.5 (half of the time) is typically referred to as the *perceptual threshold*. When the stimulus is only presented for a very short amount of time (typically less than 100 milliseconds), the intensity threshold changes systematically with presentation time (Bloch's law: higher threshold for shorter presentation time).

The idea behind why the same stimulus can sometimes be detected and sometimes be missed is that the internal representation of the stimulus is noisy and therefore fluctuates across stimulus presentations. It is assumed that the strength of this internal representation has to exceed a decision criterion or threshold for the stimulus to be detected. Note that this decision threshold, linked to neural activity, is different from the perceptual threshold, linked to stimulus intensity, mentioned earlier. For example, we might present a visual stimulus with an intensity slightly above the perceptual threshold multiple times. Note that the stimulus intensity always exceeds the perceptual threshold. Nonetheless, the stimulus will sometimes be detected (seen) and will sometimes be missed (not seen). This is because the neural activity constituting the internal representation of the stimulus fluctuates from presentation to presentation. Sometimes it exceeds the decision threshold and sometimes it does not. This concept has been formalized in *signal detection theory*.

Choices between two possible categories (e.g., "Was the object moving to the left or to the right?") are more accurate when the stimulus is less ambiguous. This is usually the case when the stimulus is of high quality, meaning that it provides a lot of information about its contents and the noise level is low. Experts refer to this situation as a high signal-to-noise ratio (SNR). A clear, focused image would be an example. When a comparison between two stimuli is required (e.g., "Was the second tone louder or softer than the first one?"), subjects are more accurate the more different the two stimuli are with respect to the relevant dimension or attribute (in this example, the sound pressure level). Furthermore, this difference typically has to be a certain fraction of the base value to achieve a particular level of accuracy. For example, when presenting a 1-kilohertz (kHz) tone and asking by how much the frequency of a second tone has to be increased to be able to detect an increase in pitch, you will probably come up with a "just noticeable difference" (JND) of approximately 5 hertz (Hz). When presenting a 4-kHz tone and asking the same question, you will probably come up with a JND of approximately 20 Hz. The ratio between the JND and the base value (in this example, 5 Hz/1,000 Hz = 20 Hz/4,000 Hz = 0.005) is referred to as the Weber fraction and, according to Weber's law (Weber-Fechner law), remains constant over a wide range of base values.

Response Time Effects

In free response paradigms, subjects are allowed to respond whenever they are ready. For example, you are watching a movie of an animal wandering through the bushes. You can watch the movie for as long as you want to. You are told to say "cat" when you see a cat and "dog" when you see a dog. In experimental situations in which subjects are presented with a random mixture of trials with varying difficulty (or signal strength), one usually observes that both accuracy and response time (RT) vary systematically with task difficulty (high SNR = low difficulty = high accuracy = fast responses; low SNR = high difficulty = low accuracy = slow responses).

Another characteristic phenomenon is observed when additional instructions to emphasize either accuracy or speed are given: When subjects are instructed to emphasize accuracy, they tend to make more accurate choices, but responses are also slower. Thus, the increase in accuracy comes at the cost of having to spend additional time on making the decisions. When subjects are instructed to emphasize speed, they tend to make faster but less accurate choices. Thus, the increase in speed comes at the cost of reduced accuracy. This phenomenon is referred to as *speed-accuracy tradeoff*. These RT effects are captured by sequential sampling models of PDM.

It has further been observed that RT varies systematically with the number of available choice options (Hick's law). RT is approximately proportional to the logarithm of the number of choices (i.e., RT increases by a certain amount every time the number of options is doubled).

Neural Correlates

Neural correlates of PDM have been studied in humans using non-invasive techniques (such as electroencephalography [EEG], magnetoencephalography [MEG], or metabolic measures of brain activity like positron emission tomography [PET] or, in particular, functional magnetic resonance imaging [fMRI]) as well as in animal models using invasive techniques (extracellular recordings of action potentials from single or multiple neurons). Whereas fMRI provides an idea about which brain areas are active during PDM, the animal recordings, because of their much higher temporal resolution, provide additional insight into how the decision unfolds over time. Overall, these studies suggest that decision-related neural activity, depending on the PDM task, can be observed in various brain areas, including the prefrontal cortex, parietal association cortex, and premotor cortex. These results suggest that the brain uses a continuous flow of information strategy rather than strict sequential processing: Perceptual decisions are not first made by some abstract PDM module, independent of how the choice is reported, and then an appropriate action is requested once the decision has been finalized. Instead, when the decision maker has a choice between several possible known actions, the decision unfolds as a competition between simultaneous action plans. Neural activity can be measured in motor planning areas well before the choice is actually made, and a transient simultaneous activation of neurons coding for different possible actions can be observed.

The time course of neural activity recorded from the parietal cortex of monkeys performing a visual PDM task (stimulus discrimination) with goal-directed eye movements as the means of reporting the choice seems consistent with sequential sampling models: initially, if the animal does not show a choice bias, the firing rates of pools of neurons coding for the different possible actions are approximately identical. The firing rates then ramp up or down with the slope of the ramp being a systematic function of the stimulus strength. Sensory evidence in favor of a particular action leads to an increase in the firing rate of the associated neurons, evidence against a particular action to a decrease. Immediately before an action is triggered, the neurons coding for the chosen action exhibit a stereotyped activity level, consistent with the idea that a decision threshold has been reached.

Theoretical Approaches

Signal Detection Theory

When a stimulus with an intensity that is close to the perceptual threshold is presented, it is sometimes perceived and sometimes missed. A breakthrough theoretical formalization of this phenomenon was signal detection theory (SDT) in the 1960s. The basic idea behind SDT is that the internal representation of a sensory stimulus is noisy and therefore fluctuates across stimulus presentations. SDT assumes that the internal representation of a sensory signal follows one particular distribution in the absence of the stimulus (noise only distribution) and another distribution in the presence of the stimulus (signal and noise distribution). It is further assumed that the stimulus is detected if the strength of the internal representation exceeds a decision criterion. Because of the variability of the internal representation, the same stimulus can sometimes cause a neural activation that does not exceed the decision criterion (stimulus is missed) and sometimes an activation that exceeds the criterion (stimulus is detected). Likewise, this concept also explains why noise can cause an observer to detect a stimulus when no physical stimulus is actually presented (the neural activation might still exceed the decision threshold).

SDT can explain why perceptual decisions are more accurate when the signal is stronger. A higher SNR is equivalent to less overlap between the two distributions of internal signals that need to be discriminated. It is therefore easier to place a decision threshold in such a way that the largest part of one distribution is on one side of the criterion and the largest part of the other distribution is on the other side of the criterion.

Sequential Sampling Models

SDT cannot explain any of the RT phenomena because time is not part of the SDT formalism. This is where sequential sampling models of PDM come into play. The basic idea behind sequential sampling models is that, in the presence of uncertainty or noise, a decision maker can benefit from sampling multiple times from the noisy distribution of internal values (neural activity) representing the stimulus. The SNR, and therefore the accuracy of the decision, can be improved by integrating or accumulating the evidence obtained from these samples. But when should one stop sampling? The sequential sampling models ("accumulator," "integrator," or "drift-diffusion" models) usually assume that a decision is made when the accumulated evidence exceeds a certain decision criterion. Mathematically, these models are closely related to the Sequential Probability Ratio Test (SPRT), which was developed during World War II. In this framework, choosing between different categories can be thought of as a race to threshold between multiple integrators, each being associated with a particular choice option.

Sequential sampling models can explain several of the experimental observations mentioned earlier. Because the integrators accumulate both signal and noise, with the signal driving the appropriate integrator closer to the correct decision boundary and noise driving any integrator closer to any decision boundary, the higher the SNR is, the more accurate decisions are. (Accumulating noise means that the actual state of the integrator is unpredictable, but the variance of the state of the integrator constantly increases. It is therefore increasingly likely that one of the decision thresholds is crossed, just by chance; which one is unpredictable, and this phenomenon therefore interferes with making a correct decision.) Decisions are also faster when the signal is stronger because the decision criterion is reached earlier (because of a steeper slope; when trials with varying difficulty [signal strength] are presented in random order, it is assumed that the decision maker is unable to make systematic adjustments to the decision threshold based on difficulty). Sequential sampling models can also explain the speed-accuracy tradeoff. When the decision criterion is moved further away from the starting point of the accumulation (the height of the decision boundary is increased in response to an emphasis on accuracy instruction), the accumulation of noise makes it less likely that a decision threshold will be crossed and therefor the decisions will be more accurate, but the signal has to be sampled for a longer duration until the decision threshold is crossed, leading to slower decisions.

Jochen Ditterich

See also Perceptual-Motor Integration; Psychophysics: Detection; Response Time; Signal Detection Theory and Procedures

Further Readings

Churchland, A. K., Kiani, R., & Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nature Neuroscience, 11,* 693–702.

Ditterich, J. (2006). Stochastic models of decisions about motion direction: Behavior and physiology. *Neural Networks, 19,* 981–1012.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience, 9,* 467–479.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873–922.

Romo, R., & Salinas, E. (2003). Flutter discrimination: Neural codes, perception, memory and decision making. *Nature Reviews Neuroscience, 4,* 203–218.

DEPTH PERCEPTION IN PICTURES/FILM

Depth perception in general can be understood as a reconstructive process that interprets the retinal image in our eye such that a three-dimensional (3-D) object arises in our mind. Pictures and films can also provide vivid impressions of depth. This pictorial depth differs in nature. It is a constructive process of its own and presents an additional level of difficulty. Normal vision allows us to glean information about an object's shape and color as well as about such things as its spatial relations, its mass, and its potential danger. Normal vision typically reconstructs the real-world object which gives rise to the retinal image with admirable precision. This is possible because our visual system is able to resolve the many ambiguities present in the retinal image. Pictorial depth is both more confined and broader than normal depth. Figure 1 illustrates the nature of the (re)constructive processes in pictorial viewing compared with normal viewing. In normal viewing, a large number of 3-D objects would qualify as permissible reconstructions that could be made on the basis of one given retinal image. This is the case for the canvas. Pictorial depth is processed in addition.

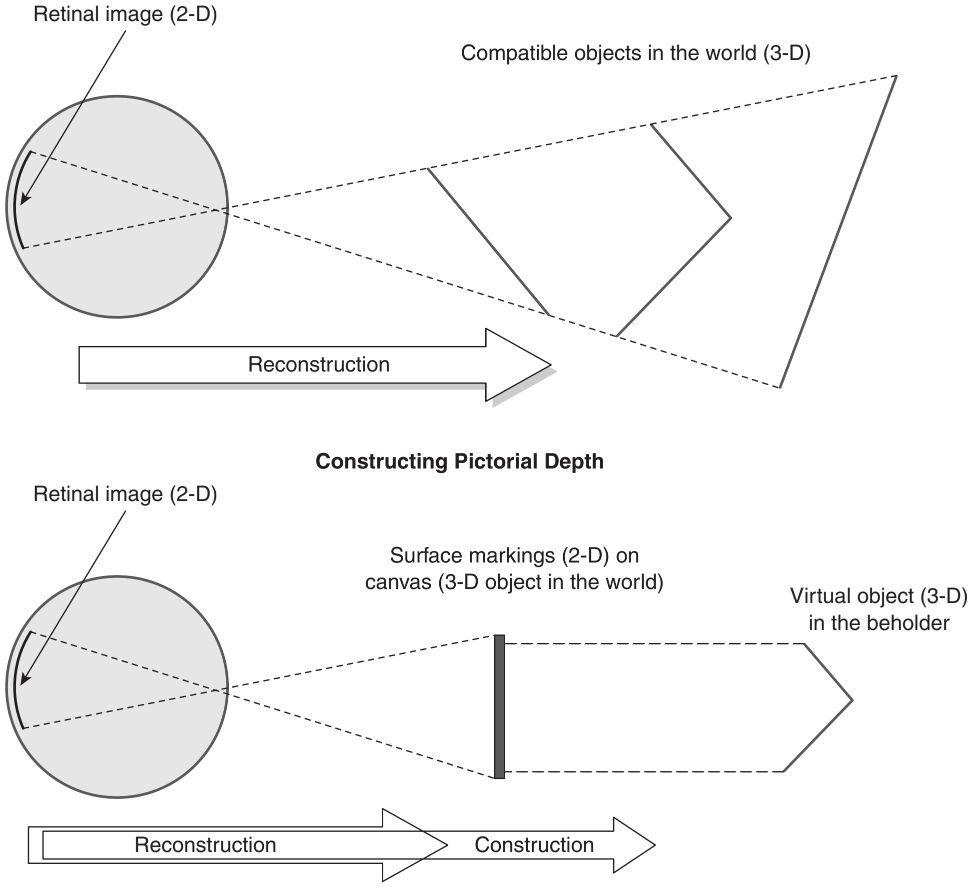


Figure 1 Reconstructing Physical Depth and the Two-Layered Process of Constructing Pictorial Depth

To date, perceptual psychologists have not been able to agree about just how the mind solves this so-called underspecification problem and singles out the one reconstruction that ends up in our awareness. The reconstruction is close to the actual 3-D object that we see, most of the time. The great physiologist Hermann von Helmholtz has argued that the visual system acts like a problem solver and uses unconscious inferences and depth cues to solve the underspecification problem. This entry describes the dual nature of pictures, the limitations of pictorial depth, and violations of spatial construction rules in frozen and dynamic renditions.

The Dual Nature of Pictures

Pictures add a new layer to the problem. To illustrate this additional layer, consider the portrait painting of Mona Lisa (or any other portrait). Assume that your vantage point is right in front of the painting and that Mona Lisa is looking straight at you. If a real person had been in her place, one side step by the observer would evade the gaze. Not so the painted portrait. You can walk several yards to either side and still, Mona Lisa is looking straight at you. Her gaze appears to follow you around, no matter where you are. This eerie quality of pictures illustrates the fundamental difference between pictorial space and normal space. Now we can take a closer look at the dual nature of picture perception, at the two layers that are both involved. On the one hand, a picture is a real-world object. It is a flat piece of canvas placed at a certain position and orientation in real space. On the other hand, the content of the picture, *Mona Lisa,* is there as well. However, she seems to obey different laws than the canvas itself does. Although the canvas does change its orientation toward us as we take a step to the side (the first layer), Mona Lisa is unmoved and does not change her orientation toward us (the second layer). The two layers of (re)construction that are present in every picture (with the exception of abstract paintings, monochromes, etc.) are at the heart of our understanding of depth in pictures.

One attempt to understand the second layer of picture perception is to conceive of paintings and photographs as seen with a disembodied eye. Although the canvas itself is seen with the normal eye, the entire content of the picture is seen as if our eye were put right in front of the canvas and remained there. Alternatively, one could think of the second layer as the content of the picture being yoked to the observer. The difference between still and moving pictures, in comparison, is rather trivial. The picture can be treated as a frozen frame of a movie. The lack of motion makes the reconstruction of the depicted object a little harder, but the process is in principle the same.

As helpful as it may be to separate the two layers of picture perception, one should not take them to be two consecutive steps. If one does, many mysteries arise, such as the mystery of the slanted picture surface. For instance, if we slant our portrait painting with respect to the observer's line of sight, the image on the retina becomes distorted. Inferring the would-be 3-D object from the painted surface is hard to begin with, and arriving at the spatial layout of the virtual object in the painting by way of the retinal image should be close to impossible when the image surface is slanted with respect to the observer. The mystery is that such a devastating effect of image slant does not occur. Its absence has long preoccupied artists, philosophers, and psychologists alike. As Maurice Pirenne has shown, the perceived aspect ratio, depth, and size of the painted object is amazingly robust. However, this robustness disappears thoroughly when we add just one layer of depiction and look at a photograph of the picture taken at an angle. Now the slanted portrait looks distorted compared with a portrait whose canvas surface was orthogonal to the camera axis.

The Limitations of Pictorial Depth

When comparing depth cues in pictures (content layer) with depth cues provided by 3-D objects, pictures are deficient in that they do not contain stereoscopic information, that is, the differences between the two views from the two eyes. The cues to actual depth provided by the degree of disparity between the two retinal images are exploited with exquisite accuracy for objects that are within our reach, and less well for objects beyond our immediate action space. In pictures, this information specifies the distance of the canvas from the observer (layer one) rather than the distance(s) specified by the content of the picture (layer two). Because the two differ, this results in conflict and renders the pictorial depth cues less powerful. It helps if we remove the disparity and regard pictures with just one eye or if we use a synopter, which duplicates the image and sends two identical views to both eyes. Gallery visitors in the 19th century were encouraged to use such devices to perceive the "true" depth in paintings. Note that viewing with one eye accomplishes the same effect by eliminating any possibility of disparity. Although this may seem paradoxical, removing the mere possibility of disparity does increase the impression of depth.

In analogy, with special viewing devices, stereo photographs can be presented to observers such that retinal disparity is reintroduced—this time artificially congruent with pictorial depth as evoked by perspective and other pictorial cues, with striking results. Typically, however, the lack of stereo information, the limited resolution of photographs, and the viewpoint independence limit pictures and distinguish them from natural viewing. In near space the extraction of depth information from pictures is ultimately limited compared with that from an actual scene. And even if we do all we can to cure that situation, such as in virtual reality displays, near space appears to be compressed in comparison. Notwithstanding this limitation, the visual system's ability to extract depth from pictures is quite remarkable. But rather than expanding on the subtler deficiencies of pictures to specify depth when the intention is to approximate normal viewing, let us now focus on the blatant opportunities to play with depth that open when the artist's intention is no longer focused on veridicality. Pictures, still or moving, can deliberately violate the laws of natural viewing and play on the visual system's inclination to construct world-like objects even if the pictorial cues are inconsistent or violate the laws of geometry or physics.

Violations of Spatial Construction Rules in Frozen Renditions

The liberty of the artist to play with pictorial depth falls into two categories. First, the artist could merely challenge our past experience. For instance, if the artist were to paint a distorted perspective rendition of a house, its sides may no longer look parallel and the roof may be bent in strange ways. Although we have never seen such a house, it could in principle exist. Our mind constructs spatial relations that are highly unusual but in principle possible. The artist has merely violated conventional wisdom but not natural law. Even size violations may depict a possible world. Imagine a photomontage with a giant seemingly touching a dwarf. Merely shrinking a pictorial object on an empty canvas will remove it farther back in virtual space; shrinking it while perspective and occlusion cues indicate its proximity will turn the object into a dwarf. The second category, violations of natural law, is less benign. For instance, objects that are clearly too heavy to float could rest in mid-air. If such a picture is produced, the visual system attempts to reconcile the depiction with natural law. The objects are typically not seen to hover above the ground but they appear withdrawn into space toward a position where they should have been had they touched the ground. The visual system proceeds as if undisturbed by the compromise that it just produced. Yet more extreme examples are so-called impossible figures, such as the famous *Waterfall* by Maurits Escher where the water continuously flows downhill but still ends up back at the top. Although the waterfall merely violates global depth relations and thereby creates the illusion of a perpetual mobile, the impossible triangle (see Figure 2), called Penrose triangle, which is attributed to Oscar Reutersvärd, appears to violate the laws of geometry and of space itself. Note that three-dimensional objects have been designed that do appear as Penrose triangles, but they work only if a stationary observer views the object from one specific vantage point, whereas the pictured Penrose triangle implies an object that could be viewed from any vantage point; such an object cannot, however, exist in reality. The power of such pictorial art is tremendous. It reveals significant aspects of the perceptual process proper, and a fundamental category of perception—space itself—is being questioned when viewing an impossible object. Spatial relations can no longer be judged, let alone in a quantitative fashion. At the same time, we could make estimates about the object's general shape (a triangle) and maybe even its size. The normal process of gleaning metric information and (re)constructing a world object is still engaged. All the counterfactual aspects of the Penrose figure do not suffice to halt the constructive perceptual process. After all, the markings on the page could be seen as a two-dimensional design. Instead, a triangle is constructed and accompanied by an emotion of awe or the feeling of being tricked.

Violations of Construction Rules in Dynamic Renditions

Moving images open another realm of possible violations of space and time. One could even say that the defining feature of movies consists in violations of normal viewing. The alter ego of the movie can levitate, walk backward without looking in that direction, fly, or shrink to fit into a keyhole. These abilities redefine space and time internally for the observer. Teleportation becomes possible and the narrative time-line can move forward or backward. Given that the constraints of normal space and time perception can be and are typically violated in motion pictures, quantitative models of space and time perception in pictures appear impossible. Both normal perception and its filmic violation must rely on a small set of common principles of depth perception. Some are almost never violated, such as occlusion relationships, but others are regularly violated, such as the laws of

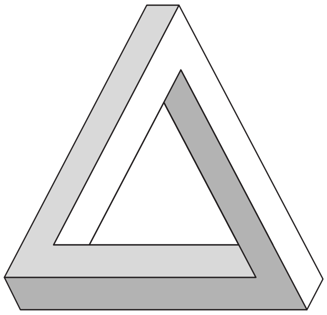


Figure 2 Impossible Figure: Penrose Triangle

perspective transformation and the apparent foreshortening of objects when different viewpoints are assumed. Here the visual system is rather tolerant, as we can see because it often fails to distinguish between the mere magnification caused by zooming-in and the more complex enlargement caused by a camera approach. The zoom has no equivalent optical transformation in natural viewing.

When it comes to optical specification of the external world, strange places can be conjured where, for example, people walk on surfaces that specify a liquid state, objects fall up instead of down, and debris assembles itself into objects. The visual system nonetheless attempts to assign specific depth values to these events and sort them into the preexisting categories of causality. In the fantasy of the virtual world to which their existence is limited, the spatial assignments cannot be verified and remain inconsequential for action. Interestingly, despite artistic freedom and the possibilities of special effects studios, the conventions of virtual space that have established themselves bear witness of a rather restrained use of the potential for violating the rules of normal viewing.

In sum, depth perception in pictures and film reveals the propensity of the visual system to construct a consistent 3-D space on the basis of optical information even if it defies the laws of physics or geometry. With regard to the limitations of pictures and movies in comparison with natural viewing, depth information from pictures is surprisingly robust.

Heiko H. Hecht

See also Film (Cinema) Perception; Pictorial Depiction and Perception; Spatial Layout Perception, Psychophysical

Further Readings

Gibson, J. J. (1971). The information available in pictures. *Leonardo, 4,* 27–35.

Goldstein, E. B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance 13,* 256–266.

Hagen, M. A. (1980). *The perception of pictures: Vol. 1. Alberti's window: The projective model of pictorial information.* New York: Academic Press.

Hecht, H., Schwartz, R., & Atherton, M. (2003). *Looking into pictures: An interdisciplinary approach to pictorial space.* Cambridge: MIT Press.

Kubovy, M. (1986). *The psychology of perspective and renaissance art.* Cambridge, UK: Cambridge University Press.

Münsterberg, H. (1916). *The photoplay: A psychological study.* New York: D. Appleton.

Peterson, M. E., Gillam, B., & Sedgwick, H. A. (Eds.). (2007). *In the mind's eye: Julian Hochberg on the perception of pictures, film, and the world.* New York: Oxford University Press.

Pirenne, M. H. (1970). *Optics, painting, and photography.* Cambridge, UK: Cambridge University Press.

DIGITAL IMAGING

In 1969, the *Apollo 11* astronauts landed on the moon, the Woodstock Festival attracted nearly 500,000 concertgoers, and George Smith and Willard Boyle invented the first charge coupled device (CCD). Although greeted with less fanfare than the other events of 1969, the CCD gave rise to *digital imaging*—the process of capturing and representing an image in a format readable by a computer. By replacing the traditional film with an electronic CCD sensor, digital imaging revolutionized photography ranging from the Hubble telescope in outer space to the cell phone in your pocket. This entry describes the capture, storage, and manipulation of digital images.

Formation

The CCD sensor is the "film" of a digital camera. It consists of a two-dimensional array of photoelectric elements that become electrically charged when exposed to light. Because the amount of charge is proportional to the light's intensity, the CCD's electrical pattern is a faithful representation of the light pattern striking the sensor. Although exquisitely sensitive to light intensity, a CCD element does not differentiate light wavelength (color). To record a color image, a color filter array (CFA) is overlaid on the CCD (see Figure 1; see also color insert, Figure 22). With a CFA, each CCD element records a limited range of wavelengths, corresponding to red, green, or blue.

The pattern of electrical charge on the CCD is transferred to the camera memory, where it is represented as an array of picture elements or pixels. A six mega-pixel digital camera, for example, corresponds to a CCD sensor with 6 million elements and a digital image with 6 million pixels. Associated with each pixel is a number between 0 and 255, where 0 corresponds to the minimum charge and 255 to the maximum charge, yielding 256 distinct intensity values. In a full resolution color image, each pixel is assigned three numbers, one for the intensity of red, one for green, and one for blue. But as noted previously, the color filter array limits each sensor element to one color and so initially each pixel represents either red, green, or blue (RGB). For a full resolution RGB color image, the camera must interpolate across neighboring pixels to fill in the missing color values (Figure 1; see also color insert, Figure 22). These three values are sufficient to give rise to tens of millions of colors. For example, red is given by the RGB triple [255,0,0], green by [0,255,0], blue by [0,0,255], yellow by [255,255,0], a dark red by [128,0,0], black by [0,0,0], and white by [255,255,255].

Three values are sufficient to represent millions of colors because a digital camera encodes wavelength in a way that is similar to the human eye. The three filters in the CFA are analogous to the three cones in the eye that selectively absorb light in either the red, green, or blue range of the color spectrum. A digital camera and the human eye invite other comparisons. For example, in the center of the visual field (the fovea), the human eye has an estimated spatial resolution equivalent to tens of millions of pixels, and a dynamic range equivalent to several thousand distinct intensity values. A dynamic range equivalent to millions of intensity values is possible for the human eye because the eye can dynamically adjust its lightness sensitivity range. Similar effects can be achieved digitally through high dynamic range (HDR) imaging in which multiple images photographed at different exposures are combined.

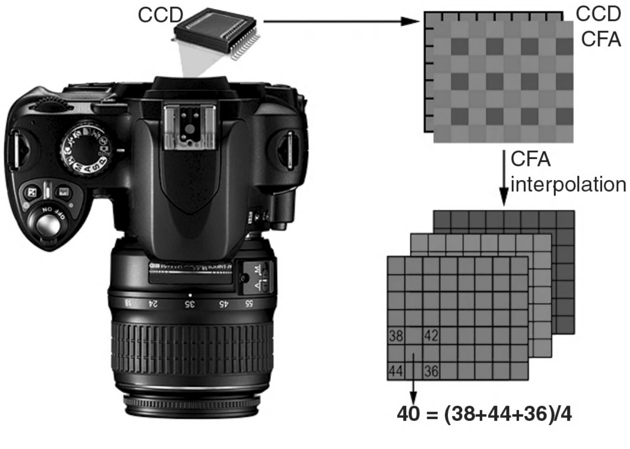


Figure 1 Color Filter Array (CFA) Overlaid onto a Charge Coupled Device (CCD) Sensor

Notes: Each CCD element records a limited range of wavelengths, corresponding to either red (gray), green (light gray), or blue (dark gray). A full three-channel RGB color image is created by interpolating the missing color pixels by, for example, averaging the recorded values. See also color insert Figure 22.

Format

A digital image in the camera and computer memory can be stored in many different formats. The most basic format, raw file, stores the pixel values directly recorded by the CCD before CFA interpolation. This format affords efficient storage because only one number is stored for each pixel, but it requires any subsequent photo editing software to perform the CFA interpolation. The remaining image formats fall into one of two categories: non-lossy and lossy. Non-lossy formats such as TIFF, PNG, and BMP, store the full RGB digital image without any loss of information. The benefit of these formats is that they permit high-quality reproductions, but with the drawback that they require significant amounts of memory for storage. The lossy GIF image format achieves compression by limiting the total number of colors in the image from tens of millions to, typically, a few hundred. The image format JPEG, perhaps the most popular lossy format, achieves compression by removing some color and high spatial frequency information (i.e., image details). These lossy image formats reduce the disk space required to store an image, but at the cost of degrading the quality of the image.

Quality

Camera manufacturers typically tout the number of pixels as the single measure of image quality. Many factors, however, contribute to the final quality of a digital image, including the camera lens, the CCD sensor, the CFA filter, and the image compression level. Low-quality lenses can lead to chromatic aberrations where the color channels are misaligned relative to one another. Lenses may also produce geometric distortions that cause straight lines to appear curved. Low-quality CCD sensors can lead to high levels of noise or graininess in the image. Low-quality CFA filters and interpolation can lead to poor color reproduction. And high levels of lossy compression can lead to significant loss of color and details. These degradations can range from the slightly to highly visible and can be particularly problematic when enlarging a digital image for printing.

In high-end digital cameras, these artifacts have been largely eliminated with results that rival those of the best traditional film cameras. In low- to mid-range digital cameras, today's six- to eight-megapixel cameras (millions of pixels) are of sufficient resolution for most consumer needs (for example, at a high resolution of 300 dots per inch [dpi], a six-megapixel digital image can be printed to a size of 8 × 10 inches).

Digital Image Processing

Before the digital revolution, the enhancement and manipulation of images required talented artists and technicians to spend long hours in the darkroom. With the advent of digital imaging, such alterations are now only a few computer mouse clicks away.

Some basic image manipulations involve systematically remapping pixel values using a lookup table (LUT). This remapping is represented graphically in Figure 2, which shows how a LUT can be used to alter the brightness of an image. The horizontal axis of the LUT corresponds to the original intensity value (in the range 0 to 255), and the vertical axis corresponds to the remapped intensity value (also in the range 0 to 255). By sliding the LUT down, for example, all pixel values between 0 and 80 are mapped to 0, and the remaining pixel values between 81 and 255 are linearly mapped into the range 1 to 175 (resulting in a darkening of the image—see Figure 2b). Functionally, consider a single pixel whose value is in the range 0 to 255. Find this value on the horizontal axis and draw a line straight up to intersect the LUT. Now draw a straight line leftward to the vertical axis to determine the new value for the pixel. This process is effectively repeated for each pixel to uniformly change the appearance of an image. Also shown in Figure 2(b) is the intensity histogram after the brightness adjustment. The horizontal axis of the histogram corresponds to the intensity value, in the range 0 to 255, and the vertical axis corresponds to the number of pixels that have the specified gray value. Darkening corresponds to a leftward shift of the histogram. Similar to darkening, brightening is achieved by shifting the LUT upward, which corresponds to a rightward shift of the histogram (Figure 2c).

As shown in Figure 2(c), image contrast can be increased by increasing the slope of the LUT. Similarly, the contrast can be decreased by decreasing the slope of the LUT. Note that both brightness and contrast enhancement can saturate pixels, where a range of values is collapsed to a single value. As a result, these operations are not reversible because once the values are mapped en masse to 0 or 255, their original values are lost. This saturation can be seen in the histogram where there is an increase in the number of pixels in the bin at 0 (black) and at 255 (white) corresponding to all of the pixels that were mapped to these values. To increase contrast while avoiding saturation, an image can be auto-scaled by adjusting the LUT such that the smallest pixel value in the image is mapped to 0 and the largest pixel value is mapped to 255.

Image contrast can also be increased or decreased through a nonlinear LUT. Shown in Figure 2(d) is the result of one such contrast enhancement, gamma correction, where each pixel value, divided by 255, is raised to the power 2, and then re-multiplied by 255 (the scaling by 255 is so that the pixel values of the gamma corrected image remain in the range 0 to 255). Varying amounts of gamma correction are often automatically applied by digital cameras and in computer monitors used to view digital images. Unlike brightness and contrast enhancement, gamma correction does not saturate any pixels. Lastly, shown in Figure 2(e) is the result of quantizing a digital image to only two gray values (also known as thresholding). The corresponding LUT

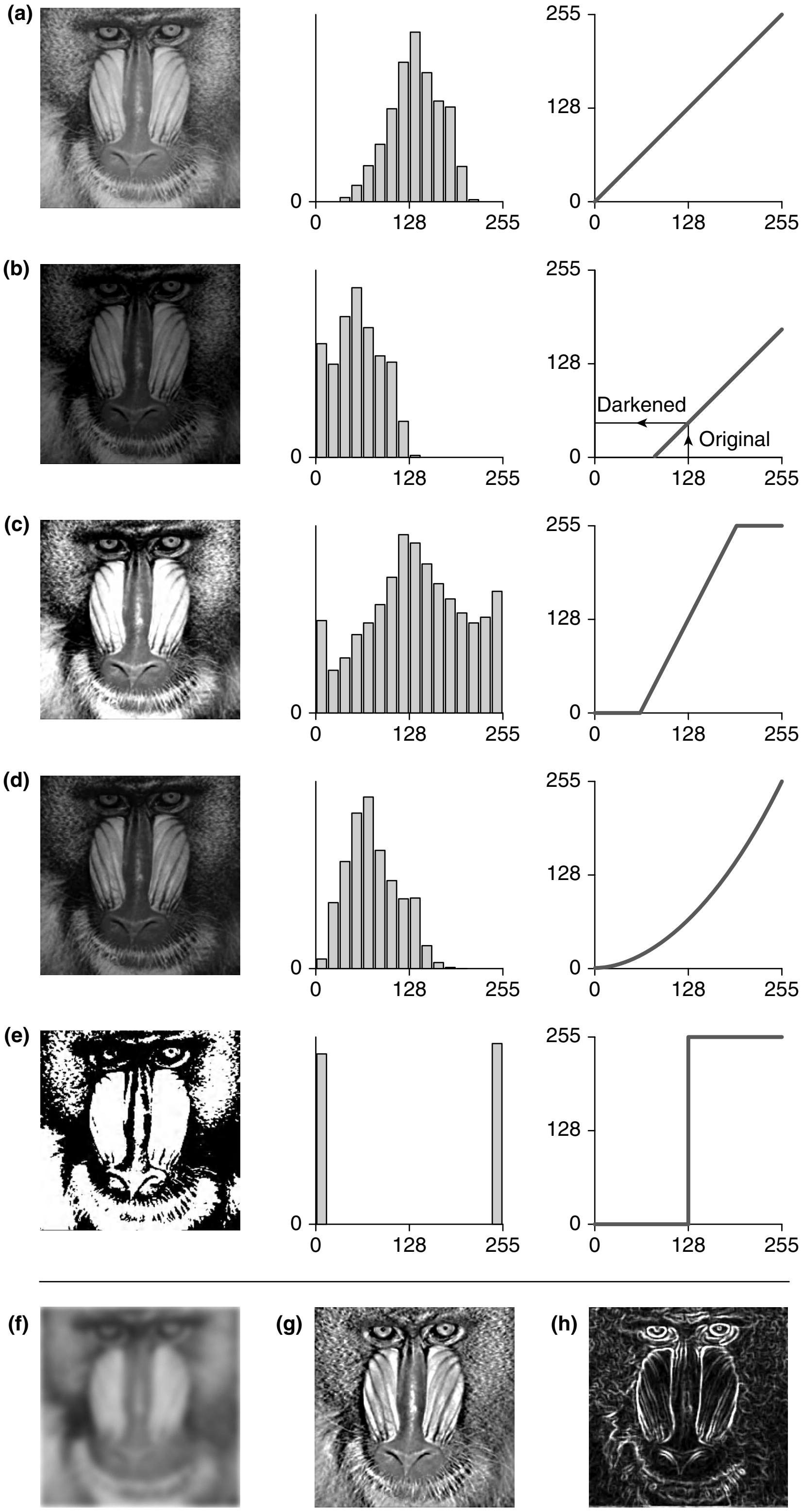


Figure 2 Basic Image Manipulations

Notes: Shown is (a) an original grayscale image that has been: (b) darkened, (c) contrast enhanced, (d) gamma corrected, (e) quantized (thresholded), (f) blurred, (g) sharpened, and (h) edge extracted. Shown in the second column of panels (a)–(e) is the corresponding intensity histogram, and in the third column is the lookup table (LUT) that embodies the transformation.

is a step function that uniformly collapses a range of intensity values to a single value.

In addition to basic intensity manipulations, digital filtering can remove or enhance information in an image. Shown in Figure 2(f) is an image that has been blurred to remove image details. Shown in Figure 2(g) is an image that has been sharpened to enhance image details. And shown in Figure 2(h) is the result of filtering that isolates edges. Each of these filtering operations involves adjusting pixel values based on the values of neighboring pixels. An image is blurred by replacing each pixel value by the average of its neighboring pixel values—the larger the neighborhood, the greater the blurring. An image is sharpened by replacing each pixel with the difference between itself and its neighboring pixels. And, an image's edges are enhanced by replacing each pixel with a combination of horizontal and vertical differences in pixel values.

Many photo-editing software packages perform these LUT and filter operations, along with many more sophisticated manipulations: red-eye can be removed from flash photographs, objects or people can be added or removed from an image, several images can be stitched together to create a wide-angle mosaic, and much more.

Implications

Advances in digital and computer technology have revolutionized photography in ways that could not have been imagined when the CCD was first invented. These advances have had an exciting impact on science, medicine, and art. These advances have also had the problematic consequence that they make it easy to doctor images in ways that are difficult to detect. Doctored images are appearing with a growing frequency and sophistication in tabloid magazines, fashion magazines, mainstream media outlets, scientific journals, political campaigns, courts of law, and our e-mail in-boxes. As a result, our trust in photography has been diminished. This in turn has given rise to the field of digital forensics and the development of techniques to authenticate digital images. In the coming years, digital and computer technology will continue to develop, and the field of digital forensics will try to help us regain our trust in photography. Although it is impossible to imagine precisely how these fields will develop, digital imaging has forever changed our relationship with visual imagery.

Hany Farid

See also Computer Vision; Eye: Structure and Optics

Further Readings

Farid, H. (2008). Digital Doctoring: Can we trust photographs? In B. Harrington (Ed.), *Deception: Methods, motives, contexts and consequences.* Stanford, CA: Stanford University Press.

Gonzalez, R. C., & Woods, R. E. (2008). *Digital image processing* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Russ, J. C. (2006). *The image processing handbook* (5th ed.). Boca Raton, FL: CRC Press.

DIGITAL PHOTOGRAPHY

See Digital Imaging

DIRECT PERCEPTION

When we open our eyes, it seems to us that we are aware of ordinary objects and events such as tomatoes, elephants, and bouncing balls. *Direct perception* means that we are indeed aware of such environmental objects, without mediating representations or inference processes. Yet this appears to present a paradox, for we are not in physical contact with environmental objects, but only with the light that arrives at our receptors, the neural signals it elicits, or resulting sensory states. The problem of perception is how it is possible to be aware of environmental objects and events when we are only in immediate contact with such intervening states.

There are three general responses to the paradox. The *phenomenalist* concedes that we are only aware of sensory phenomena, not environmental objects. When we look at a tomato, we have sensory impressions of roundness and redness, but we do not perceive the tomato itself. For all we can tell, we could be living in the virtual world of *The*

Matrix, with an illusory tomato induced by electrical stimulation of our brains.

The *indirect realist* goes a step further, arguing that the perceiver can at least make inferences about environmental objects based on sensory states. When we look at a tomato, we construct a mental representation of it; we are directly aware of this internal representation, and indirectly perceive the tomato.

The *direct realist* is willing to go all the way, holding that the perceiver is directly aware of environmental objects, not just their mental proxies. When we look at the apocryphal tomato, we are in perceptual contact with and have access to the tomato itself, without mediating inferences or representations. The challenge for the direct view is to articulate a theory of perception that can sustain this claim.

The position one takes on this philosophical question has practical consequences for psychology and cognitive science. Most researchers have adhered to an indirect view, ranging from the constructivist theories of the 1950s through 1960s, in which perception is an internal construction based on impoverished cues and rules of inference; to the computational approach of the 1970s through 1980s, in which perception is the computation of an internal representation based on insufficient input and prior assumptions; and the Bayesian framework of the 1990s through the early 2000s, in which perception is statistical inference based on ambiguous image data and prior knowledge of probabilities.

Yet, during the last century, prevailing opinion among philosophers of mind has shifted from indirect theories to some version of direct perception. This entry briefly presents several of the arguments for and against direct perception and introduces the main scientific theory that aims to sustain a direct view, the ecological approach developed by James J. Gibson and his followers.

Problems for the Indirect View

Arguments for direct perception often begin with a concern that the indirect view is insupportable, easily sliding into phenomenalism and even skepticism. On the indirect view, perception consists of getting a representation of the world into the observer's head; the direct realist's concern is that this doesn't really solve the problem of perception.

Representationalism

Traditionally, the indirect view has claimed that the perceiver is aware of an external object in virtue of being aware of something else, such as an internal mental or brain state. These intervening states may be representations, mental images, sensory impressions, sense-data, neural signals, or retinal images. However, this view commits what is called the *representationalist fallacy.* The claim that the perceiver is aware of an internal state confuses the *object of awareness* (the actual tomato) with the *vehicle of awareness* (the intervening state). In particular, to say that the perceiver is aware of an internal representation sets up a logical regress, implying an inner perceiver (homunculus) who must create an inner representation of the representation. The traditional version of the indirect view is thus confused.

The regress can be avoided by claiming that awareness of an external object is *constituted* by having an internal representation of it; the perceiver is not aware of the representation itself, but of what it represents (its content). Thus, one perceives the tomato (the object of awareness) in virtue of *possessing* an internal representation of it (the vehicle of awareness). This move satisfies some philosophers that perception is direct in the traditional sense, yet on this view, the perceiver experiences the content of a representation rather than the living tomato. The representation must somehow be derived from the visual input by a process that establishes its content. This leads to a skeptical argument known as Hume's problem.

Hume's Problem

Hume's problem, posed by David Hume, also called the semantic problem or grounding problem, is the problem of how internal representations get connected to the world. If perceptual awareness consists of having representations, how does the perceptual system determine the environmental entities to which they correspond? Without some independent, extrasensory access to the world, there appears to be no way to establish which internal states indicate which environmental properties, or which representations stand for tomatoes and which for elephants. The perceiver is trapped in a closed universe of sensory phenomena or uninterpretable representations.

The indirect solution is *inference to the best explanation:* The perceptual system infers a representation of the world that best accounts for the order in sensory input. For example, a particular sequence of gray blobs with an extended protuberance may be best explained by the presence of an elephant, rather than a tomato. However, as Hermann von Helmholtz understood by the mid-19th century, this inference process presumes that the perceptual system already possesses knowledge about (1) the structure of the world, including the sorts of entities that exist and predicates to describe them, and (2) how the world structures sensory input, such as a theory of image formation and transduction.

The trouble is that such prior knowledge must somehow be acquired, again in an extrasensory manner. As Gibson observed, knowledge of the world cannot be explained by supposing that knowledge of the world already exists. A common response is that prior knowledge has evolved via natural selection or learning, but as pointed out by Michael Turvey and Robert Shaw, this seems to require an organism that already has a working perceptual system—including the requisite prior knowledge—as a precondition. The indirect position thus appears to be circular.

Inference

There is a further problem with treating perception as a process of inference. Inference is a logical relation that holds between conscious mental states (beliefs, thoughts, statements) corresponding to premises and conclusions. But as we have just seen, if we are to avoid the representationalist fallacy, perception cannot be based on conscious awareness of internal states.

If the perceptual process is unconscious, then whatever else it may be it cannot be inferential; the same goes for related terms such as *hypothesis, clue, evidence,* and *assumption.* The notion of perception as *unconscious inference,* originally suggested by Helmholtz, is thus inconsistent. Computational theories seek to avoid this objection by treating perception as a process of computation over representations, but this leaves Hume's problem unresolved.

Obstacles for the Direct View

These objections to indirect perception tend to drive theorists either to outright phenomenalism or some form of direct perception. But the direct view is not without its own obstacles.

The Argument From Illusion or Hallucination

Direct perception seems to imply that perception is veridical, that is, that the world is as it appears to be. Yet in cases of visual illusions and hallucinations, the world is *not* actually as it appears to be. According to the argument from illusion or hallucination, such cases undermine the claim that perception is direct:

1. Assume it is possible to experience illusions or hallucinations that are subjectively indistinguishable from cases of veridical perception. For example, we may have an alcohol-fueled experience of a dancing pink elephant that seems compellingly real.
2. In such cases, we are aware of *something,* but it is not a normal physical object. Hence, we are aware of a *nonnormal* object, such as a dancing pink sense-datum or an elephantine representation.
3. If we are aware of sense-data in cases of hallucination, and we cannot distinguish them from veridical cases, this implies we are aware of sense-data in veridical perception as well. Therefore, perception is not direct. As Hippolyte Taine famously remarked, veridical perception is merely hallucination that happens to be true.

The argument gets its force from the parsimony of admitting only one kind of object of awareness, nonnormal sense-data. The *disjunctivist* response (named after the logical *or* relation) denies that there is only one kind of object of awareness, rejecting step 3: It is logically possible that we are aware of nonnormal objects in cases of hallucination, while being aware of normal objects in cases of perception. Hence, we are aware of a normal *or* nonnormal object, although this is less parsimonious.

One might go further and deny step 2, arguing that when we hallucinate a dancing pink elephant, we are not aware of any *object* at all, normal or nonnormal. We are merely having a subjective experience that arises from some neural activity in our visual systems. If we can account for hallucinatory experiences in this manner and avoid introducing a spooky class of nonnormal objects, this approach appears more parsimonious. Although an observer may not be able to tell "from inside" whether a subjective experience of an elephant is veridical or hallucinatory, "from outside" the former involves awareness of a normal environmental object whereas the latter does not.

The Argument From Underspecification

Another argument against direct perception holds that perception is underdetermined by the available information. The stimulation at the receptors is said to be inherently impoverished or ambiguous, insufficient to uniquely specify environmental objects and events. A tomato is a three-dimensional spherical object, but its retinal image is just a two-dimensional circular form; working backward, this image could correspond to a flat disk or various ellipsoidal objects stretched along the line of sight (as illustrated by Adelbert Ames' famous distorted room).

To solve this "inverse optics" problem, the perceiver must somehow "go beyond the information given," inferring or computing a representation of a spherical tomato based on prior knowledge or assumptions. Perception is consequently indirect. However, if we accept that all stimulation is ambiguous, it is not clear how a working perceptual system (with the requisite prior knowledge) could evolve in the first place, returning us to Hume's problem.

Gibson's Ecological Realism

Gibson sought to dissolve these paradoxes by reformulating the problem of perception to sustain an ecological version of direct realism. Instead of treating perception as a process of getting a representation of the world into the perceiver's head, he considered it to be a *relation* between the perceiver and its environment that puts the perceiver in epistemic (knowing) contact with aspects of the environment that are relevant to its way of life.

The starting point for the ecological approach is Gibson's principle of *animal-environment mutuality*. Because a species and its ecological niche co-evolve, the perceiver is adapted to regularities of its niche that support successful perceiving and acting. In particular, perceptual systems have evolved to detect patterns of stimulation (optical, acoustical, mechanical, chemical) that uniquely specify behaviorally relevant aspects of the environment (*affordances*). The frog's visual system, for example, is tuned to the motions of small black specks that specify edible prey and large expanding patches that specify looming threats. Because such information is specific, not equivocal, perceptual systems can evolve by adapting to the available stimulation without prior knowledge, putting organisms in direct contact with the affordances of their niche and guiding effective action. This offers a means to resolve Hume's problem.

Gibson thus rejected the argument from underspecification. Information, in Gibson's sense, is a complex spatiotemporal pattern of stimulation that is lawfully structured by the environment and is specific to environmental features, under ecological constraints. The perceptual system need not "know" or "assume" these constraints explicitly, merely adapt to the informational regularities they guarantee. Within the frog's niche, for example, moving black specks overwhelmingly correspond to edible insects (although there may be an occasional blowing bit of bark), so adapting to this motion pattern promotes successful action and survival. Research has shown that there is sufficient information available to a moving observer to specify many (but not all!) aspects of environmental layout, objects, events, their relation to the perceiver, and what they afford for action.

Direct perception does not mean there is no perceptual process or mechanism, only that the process is based on specific information and hence non-inferential. In Gibson's view, perceptual systems are active attentional systems with ascending and descending neural pathways that seek and "resonate" to information. These systems become tuned to available information through evolution and learning and come to embody an implicit mapping from patterns of stimulation to behaviorally relevant environmental features. When information is detected, the perceiver is put in direct contact with its environment: The object of awareness is the environmental situation and the vehicle of awareness is the resonating perceptual system.

Optic Flow and the Control of Locomotion

A primary example of such information is the *optic flow field*. When an observer moves through the environment, a radial pattern of motion is produced at the eye, much like the outward expanding motion on the screen of a driving video game. The focus of expansion specifies the observer's direction of travel, the rate of expansion specifies the time-to-contact with objects, and velocity gradients specify the layout of surfaces.

Species ranging from insects to primates possess neural pathways that are selective for such complex flow patterns. Research by David Lee, William H. Warren, and others has shown that optic flow (along with other information) is used to control locomotion, including balance, braking, steering, and landing and escape responses. Many behaving animals thus appear to be connected to their environments by the optic flow field.

Invariants and 3-D Shape

How we perceive three-dimensional (3-D) shape with a two-dimensional (2-D) retina is often considered a classic example of underdetermined perception. But Gibson argued that such constant properties as surface shape are specified by constant spatiotemporal patterns of stimulation, which he called *higher-order invariants*. Recent work by Jan Koenderink, James Todd, and others has begun to bear out this claim: Higher-order patterns of texture, shading, optic flow, and binocular disparity indeed specify the 3-D shape of a smooth surface. However, this information is not sufficient to determine Euclidean shape (absolute depth, slant, curvature) as researchers have typically assumed, but rather qualitative surface shape (hills, dales, ridges, valleys, and plains). These qualitative features are perceived more accurately and reliably, and are theoretically sufficient to identify objects and guide actions like grasping.

An important implication of this research is that whether we consider perception to be veridical depends on the properties we assume are perceived. Perception can be accurate and direct only insofar as the property is specified by information.

The definition with which we began can now be refined. In Gibson's view, direct perception means that the perceiver is in epistemic contact with behaviorally meaningful aspects of its environment by virtue of detecting specific information, without mediating representations, inference processes, or prior knowledge.

William H. Warren

See also Action and Vision; Constancy; Ecological Approach; Indirect Nature of Perception; Spatial Layout Perception, Psychophysical

Further Readings

Chemero, A. (2000). Anti-representationalism and the dynamical stance. *Philosophy of Science, 67,* 625–647.

Crane, T. (2005). The problem of perception. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy.* Retrieved from http://plato.stanford.edu/archives/spr2005/entries/perception-problem/

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Huemer, M. (2001). *Skepticism and the veil of perception.* Lanham, MD: Rowman & Littlefield.

Pepping, G. J., & Grealy, M. L. (Eds.). (2007). *Closing the gap: The scientific writings of David N. Lee.* Mahwah, NJ: Lawrence Erlbaum.

Shaw, R. E., Turvey, M. T., & Mace, W. M. (1981). Ecological psychology: The consequence of a commitment to realism. In W. Weimer & D. Palermo (Eds.), *Cognition and the symbolic processes II* (pp. 159–226). Hillsdale, NJ: Lawrence Erlbaum.

Todd, J. T. (2004). The visual perception of 3D shape. *Trends in Cognitive Sciences, 8,* 115–121.

Warren, W. H. (2005). Direct perception: The view from here. *Philosophical Topics, 33,* 335–361.

Warren, W. H. (2008). Optic flow. In A. I. Basbaum, A. Kaneko, G. M. Shepherd, & G. Westheimer (Eds.), *The senses—a comprehensive reference: Vision II* (Vol. 2, pp. 219–230, T. D. Albright & R. Masland, Eds.). Oxford, UK: Academic Press.

DIRECT REALISM

See Naïve Realism

DYSLEXIA

Dyslexia is marked by unusually slow reading development in children who are otherwise normally developing. Although most children learn to read with little conscious effort, children with dyslexia lag well behind their peers in acquiring this fundamental skill. These children fail to learn to read adequately despite normal hearing, cognitive, neurological, and emotional abilities. The disorder is not uncommon, with an estimated 5 to 15% of all children having significant difficulties in learning to read. Despite early reports, there is no sex link, with an equal preponderance in boys and girls. The term *dyslexia* is occasionally also used to refer to individuals with reading disturbances following stroke, although the more common term in that case is *alexia,* which is not discussed in this entry.

Diagnosis of dyslexia usually precludes general sensory impairments such as severe loss of hearing or vision, or a broader cognitive deficit. Affected individuals show slow laborious reading marked by errors in single word identification and poor text comprehension. In addition, they tend to show poor generalization, such that they have difficulty reading unfamiliar words or nonsense words, indicating that they have failed to learn consistencies in the relationship between spelling and sound.

There is a relatively broad consensus that dyslexia stems from a deficit in processing phonology, or the basic sounds of words. Phonological abilities seem to play an important role in successful reading development because they are an inherent part of reading: Recognizing a word involves translating orthography (a word's visual form) into a phonological code (the sounds that the word is composed of). Thus, the phonological deficit hypothesis of dyslexia holds that affected individuals have difficulty learning to read because of a deficit in how they process phonological information. This does not necessarily imply a perceptual deficit, as is discussed further in this entry.

Consistent with this, affected individuals have difficulty with *phonological awareness.* Measures of this include *rhyme awareness* (judging whether two words rhyme or producing words that rhyme with a sample word); *phoneme blending* (assembling a set of phonemes into a single word); *phoneme counting/tapping* (counting or tapping a finger to indicate distinct phonemes in a word); and *phoneme deletion/elision* (repeating a word after removing one or more of its constituent sounds). Importantly, phonological awareness relates to processing spoken words, and performing such tasks does not hinge on the ability to read words. Listeners must understand that auditory words are not wholes, but instead can be broken down into constituent sounds.

Phonological awareness allows the beginning reader to analyze spoken words into phonemes, which can in turn be associated with orthography. It also allows the beginning reader to make analogies among similar sounding words, which typically share similar spellings, allowing beginning readers to learn new words quickly. Phonological difficulties in dyslexia are thus hypothesized to make it more difficult to process or identify the similarities and differences among sounds, ultimately making learning to read more arduous. Some of the earliest evidence supporting this view comes from the finding that phonological skills in pre-reading children are highly predictive of later reading achievement. Thus, preliterate children with phonological difficulties tend to develop reading deficits later on. Furthermore, phonological deficits tend to persist into adulthood, even in cases where reading difficulties have been effectively remediated.

One challenge to the phonological deficit view is that, despite widespread findings of poor phonological abilities in dyslexia, it has been difficult to identify what specific aspects of phonology are affected, and what their root cause might be. The opacity of the underlying deficits in dyslexia may result from the nature of the tasks used to measure them. These tasks are typically metalinguistic, requiring individuals to provide overt judgments about phonology. For instance, although rhyme awareness early on is a good predictor of eventual reading success, by ages 8 or 9, rhyme awareness becomes too easy to adequately differentiate good and poor readers. Metalinguistic tasks are also influenced by nonverbal processes such as attention and working memory. Thus, although children with dyslexia have poor phonological awareness, performance on these tasks may be constrained by age, reading skill, and other compensatory strategies that might obscure the cognitive bases of an underlying phonological deficit.

One approach to better refining our understanding of this deficit comes from studies of speech perception. Typically developing children and adults exhibit categorical perception of phonemes (especially consonants), marked by stronger between-category than within-category discrimination and non-monotonic categorization curves along equally spaced acoustic continua. For instance, it is possible to manipulate the sound "d" in a way that causes listeners to perceive it as "t," by modifying a single acoustic parameter called voice onset time (VOT). Typically, VOTs 30 milliseconds (ms) or greater are perceived as "t," and VOTs less than this are perceived as "d." Interestingly, listeners are poorer at discriminating two exemplars within a category (e.g., 10 vs. 20 ms VOT) than two exemplars from different categories (e.g., 20 vs. 30 ms VOT), even when the magnitude of this acoustic difference is the same in both cases.

In contrast, children with dyslexia have shown weaker between-category discrimination and less reliable categorization of sounds at endpoints or midpoints. Likewise, these children show atypical event-related potentials (ERPs) in response to auditory speech sounds. Importantly, these measures can be obtained for unattended stimuli, suggesting that effects are not caused by differences in attention, short-term memory, or motivation in children with dyslexia versus controls.

There remains considerable disagreement concerning perceptual deficits in dyslexia. Some studies have failed to observe such deficits, and others have only observed them in subsets of dyslexic children. Thus, it remains an open question whether speech perception difficulties can adequately account for the phonological awareness deficits and ultimately reading deficits in dyslexia.

Equally contentious is whether perceptual deficits in dyslexia are speech-specific, or caused by a more general auditory processing impairment. One theory holds that these children have a rapid temporal processing (RTP) difficulty such that they cannot process rapidly changing sensory information. This deficit would lead to difficulty perceiving the types of acoustic features that are important to phoneme recognition (especially consonants). However, this theory also holds that this deficit is domain-general such that affected individuals also have difficulty with nonspeech stimuli involving rapid temporal cues (e.g., temporal order judgment, gap detection).

Finally, two alternative theories of dyslexia bear some mention here. The first is the magnocellular deficit theory, which suggests that individuals with dyslexia have a low-level visual deficit that impairs how affected individuals process visual words. Consistent with this, some studies have reported reading-impaired children have difficulties with visual contrast detection and other types of tasks linked with the magnocellular visual channel (this is the neural pathway associated with resolving coarse object detail and motion and that contrasts with the parvocellular pathway used for resolving fine detail and color). The second theory is the cerebellar deficit hypothesis, which suggests dyslexic children have abnormal cerebellar function that affects the development of automatization in speech-related function, which in turn leads to reading delay.

Amy S. Desroches and Marc F. Joanisse

See also Eye Movements and Reading; Language; Speech Perception; Word Recognition

Further Readings

Desroches, A. S., Joanisse, M. F., & Robertson, E. K. (2006). Specific phonological impairments in dyslexia revealed by eyetracking. *Cognition, 103*(3), B32–B42.

Pugh K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., et al. (2001). Neurobiological studies of reading and reading disability. *Journal of Communication Disorders, 34*(6), 479–492.

Ramus, F. (2003). Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology, 13*(2), 212–218.

Snowling, M. J. (2000). *Dyslexia* (2nd ed.). Oxford, UK: Blackwell.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45,* 2–40.

E

EAR

See Auditory Processing: Peripheral

ECHOLOCATION

Echolocation is an active perceptual system, used by some animals to localize and identify objects in complete darkness. The active component of echolocation is the animal's production of sounds that reflect from objects in the environment, and perception arises from listening to the features of these echoes. Echolocation has been documented in bats, marine mammals, some species of nocturnal birds, and in blind or blindfolded humans to a limited extent. This entry describes echolocation signals, perception by echolocation, and neural mechanisms for echo processing.

Consider an echolocating bat hunting insects in the night sky. Light levels are low, and the bat need not rely at all on its vision. Obstacles are present, but the bat dodges branches and telephone wires with ease. As the animal flies, it produces vocalizations that are in the ultrasonic range, from 20 kilohertz (kHz), the upper frequency limit of the human ear, up to 200 kHz in some species. These high-frequency signals are well suited to return echoes from small objects, such as insects, because the wavelengths of the signals are short with respect to the reflecting surfaces.

During each attack on a prey item, the bat relies on changing echo features to guide its movement. Imagine an insect located off to the right of the bat's midline when first detected. The bat computes the horizontal and vertical positions of the insect from differences in the perceived arrival time, intensity, and spectrum of echoes at the two ears. After determining the two-dimensional position of the insect, the bat aims its head at the prey, produces sonar vocalizations, receives new acoustic snapshots of the target, and redirects its flight path. Also essential to the bat's successful pursuit of prey is the estimation of target distance. The bat determines target distance from the time delay between its sonar vocalizations and the returning echoes. It uses the three-dimensional (3-D) information about target location to guide the features of its sonar vocalizations and to position itself to grasp the insect with its wing or tail membrane.

Echolocation Signals

The signal types used by different bat species vary greatly, but all contain some frequency modulated (FM) components, whose frequency varies over the duration of the signal. FM signals are well suited to carry information about the 3-D position of objects because each frequency in the broadband signal activates different populations of frequency-tuned neurons, whose onset responses can collectively mark echo arrival time at the bat's ears. Constant frequency (CF) components are sometimes combined with FM components, and these tonal sounds are well suited to carry information about target movement through Doppler shifts in the returning echoes. A Doppler shift is a change in sound frequency that is proportional to the relative velocity of the source (in this case, the sonar target) relative to the listener (in this case, the bat). If the bat approaches the target, the echo frequency increases, noted as a positive Doppler shift. Evidence indicates that bat species using both FM and CF signals show individual variation in signal structure that may facilitate identification of self-produced echoes. One echolocating bat species, *Rosettus aegyptiacus,* produces clicks with the tongue. The most widely studied echolocating marine mammal, the bottlenose dolphin (*Tursiops truncates*), produces brief clicks, typically less than 50 microseconds in duration, with spectral energy from 20 kHz to well over 150 kHz. The temporal and spectral characteristics of the dolphin clicks depend on the acoustic environment in which sounds are produced.

Perception by Echolocation

Detection of objects by echolocation depends on the strength of returning echoes. Large objects, such as trees, buildings, and the ground reflect strong echoes and can be detected at distances greater than 10 meters. Small insect prey can be detected by bats at distances of 3 to 5 meters. Once an animal using echolocation detects a target, it relies on interaural time and amplitude differences, spectrum, and arrival time of the echoes to localize the object in 3-D space. Laboratory studies of target tracking along the horizontal axis in bats suggest an accuracy of about 1 degree. Accuracy of vertical localization in bats is approximately 3 degrees. Behavioral studies of distance discrimination in bats using FM echolocation calls report thresholds of about 1 centimeter (cm), corresponding to a difference in arrival time of approximately 60 microseconds. In marine mammals, behavioral experiments show that the dolphin's horizontal angular discrimination depends on sound frequency; for pure tones, horizontal discrimination is 3.6 degrees at 6 kHz, 2.1 degrees at 20 kHz, and 3.8 degrees at 100 kHz. Localization accuracy in dolphins improves with broadband click stimuli, with thresholds in both the horizontal and vertical planes of less than 1 degree. Along the distance axis, dolphins can discriminate target range differences of approximately 1 cm, similar to that of the bat.

Many bat species that use CF echolocation sounds are specialized to detect and process frequency and amplitude modulations in returning echoes that are produced by the moving wings of flying insect prey. The CF signal components are relatively long in duration (up to 100 milliseconds), sufficient to encode echo changes that occur from the movement of a fluttering insect over one or more wing-beat cycles. For example, the greater horseshoe bat (*Rhinolophus ferrumequinum*) can discriminate frequency modulations in the returning echo of approximately 30 hertz (Hz) (less than 0.5% of the bat's 83 kHz CF signal component).

In addition, this species can discriminate fluttering insects with different echo signatures. Several bat species that use CF signals for echolocation adjust the frequency of sonar transmissions to offset Doppler shifts in the returning echoes, the magnitude of which depends on the bat's flight velocity. These adjustments in the frequency of the CF echolocation calls, referred to as Doppler shift compensation, allow the bat to isolate small amplitude and frequency modulations in sonar echoes that are produced by fluttering insects.

More complex aspects of perception by echolocation have been examined in some bat species. Free-flying bats can discriminate between small objects differing in shape and texture. 3-D target recognition of fluttering insects viewed from novel perspectives has been reported in the greater horseshoe bat.

Dolphins can use echolocation to discriminate between objects made of different materials and shapes. Target dimensions and materials influence the intensity, spectrum, and temporal characteristics of the returning echoes, providing a number of potential cues for complex target discrimination by dolphins.

Neural Mechanisms for Echo Processing

Decades of research have revealed specializations for the processing of sonar echoes in the auditory system of bats. In dolphins, studies of neural processing have been limited, but early work documents brain responses to sounds in the ultrasonic range of marine mammal echolocation calls.

Some bat species using CF echolocation calls show specializations in the peripheral (inner ear) and central (brain) auditory systems for processing echoes in the CF range of their echolocation sounds. The greater horseshoe bat, for example, adjusts the frequency of its sonar vocalizations to receive echoes at a reference frequency of approximately 83 kHz. Its auditory system shows a large proportion of neurons devoted to processing this reference frequency, and an expanded representation of 83 kHz can be traced to mechanical specializations of this bat's inner ear.

Other specializations in the bat central auditory system for processing echolocation signals that may play a role in the perception of target distance. In several bat species, researchers have identified neurons in the midbrain, thalamus, and cortex that respond selectively to pairs of FM sounds, separated by a particular delay. The pairs of FM sounds simulate the bat's sonar vocalizations and the returning echoes, and the time delay separating the two signals corresponds to a particular target distance. The pulse-echo delay that elicits the largest facilitated response is topographically organized in some bat species that use CF signal components. Most delay-tuned neurons in the bat auditory system respond in the range of 2 to 40 milliseconds, corresponding to target distances of approximately 34 to 680 cm. These best delays represent a biologically relevant range for localizing prey items using echolocation.

Cynthia F. Moss

See also Animal Depth Perception; Animal Frequency and Pitch Perception; Depth Perception in Pictures/Film

Further Readings

Griffin, D. (1958). *Listening in the dark*. New Haven, CT: Yale University Press.

Moss, C. F., & Schnitzler, H.-U. (1995). Behavioral studies of auditory information processing. In R. R. Fay & A. N. Popper (Eds.), *Springer handbook of auditory research: Hearing by bats* (pp. 87–145). Berlin: Springer.

Suga, N. (1990). Biosonar and neural computation in bats. *Scientific American, 262*(6), 60–68.

Thomas, J. T., Moss, C. F., & Vater, M. (Eds.). (2004). *Echolocation in bats and dolphins*. Chicago: University of Chicago Press.

ECOLOGICAL APPROACH

Ecological psychology, as it applies to the domains of perception and perceptual development, refers to the perspective developed during a 30-year period by J. J. Gibson and Eleanor J. Gibson, and elaborated during the past two decades by their students and colleagues. The impetus for this approach grew largely from J. J. Gibson's work early in his career on several practical problems, such as how individuals control their movement when driving a car or landing a plane. These investigations led him to conclude that traditional theories of perception, and their supporting evidence mostly gathered in laboratory conditions, did not apply well to more everyday circumstances of perceiving. This entry describes major concepts in ecological optics, perceptual learning and development, and philosophical implications of ecological realism.

The adjective *ecological* in ecological psychology refers to two distinguishing characteristics of Gibson's approach to perception that sets it apart from more traditional theories.

1. Traditional theories of visual perception begin their analysis with consideration of patterns of stimulation on the sensory receptors (i.e., the retina), and the ensuing patterns of neural firing to sites in the brain. The starting point for an ecological approach to perception is an analysis of the environment within which a species has evolved (i.e., its *econiche*). The environment for terrestrial organisms is filled with *substances* of a wide-ranging variety, from soils and grasses to bodies of animals and water. In the case of visual perception, these substances can be perceived because of the way light from a radiant source (e.g., the sun) interacts with their *surface* properties (e.g., orientation to light source, texture, pigmentation). The resulting array of reflected light that has been structured by such surface properties fills the *medium* (the air), and this array of structured light surrounding the individual (*ambient optic array*) can function as potential *information* for perceiving. The study of how light in the medium is structured by surfaces is referred to as *ecological optics*. J. J. Gibson proposed that psychologists begin to consider higher-order structure in the ambient optic array of information for perceiving.

2. Traditional theories of visual perception take as their primary focus a stationary perceiver positioned at a fixed observation point. In contrast, the ecological approach takes as its primary focus the *dynamic perceiver-environment relationship*. What is dynamic about the perceiver-environment relationship? On the one hand, environmental conditions are not static, but change over various units of time (e.g., seconds, hours, days, seasons) requiring perceivers to keep abreast of conditions. On the other hand—and significantly—perceiving involves ongoing, exploratory actions of the individual (the functioning of *perceptual systems*) in the detection of stimulus information. Just as it is far easier to identify an object through active touch (manipulating the object in one's hand) than passively grasping it, perceiving through vision is facilitated by actions of the individual, such as moving the head and body in relation to objects and to the overall environmental layout.

Why do actions of the body contribute to detection of stimulus information? They produce perceived changes in the ambient optic array, and in so doing reveal that which does not change, what is invariant in the array of reflected light. (An invariant is a set of relationships among structures that do not change across transformations of those structures over time.) Invariance in the ambient array is posited to be specific to stable and persisting features of the environment. For example, if a perceiver walks around a table, certain relationships, such as the adjacent order of corners and edges, will remain invariant across successive views, displaying the specific rigid structure of this object. Actions of the various perceptual systems play an essential role in revealing invariants in the ambient array over time. Concurrently, the perceived changes in the ambient array that are generated by the perceiver's actions provide information for self-movement.

Major Concepts in Ecological Optics

Texture Gradients

Early work in ecological optics emphasized the perception of filled space rather than abstract space. J. J. Gibson proposed that "space" could not be perceived in the absence of perceiving a continuous background surface and that the topic of "space" perception would be more accurate if called "ground" perception. The texture of the ground (e.g., dirt and clumps of grass) projects a regular pattern of decreasing image size, and increasing density, from observer to horizon. This regular change is an example of a *texture gradient*. The texture gradients of surfaces, especially the ground stretching to the horizon, became the basis of reformulating numerous classical problems, two of which are distance perception and size perception. The perception of the texture gradient of a surface from observer to horizon establishes a scale of perceivable distance for a given height of the observer's eyes. The topic of size perception often is stated in a way that features the problem of *size constancy*. This asks, How is it possible to see an object as having a constant size with changes of viewing distance when the image size changes with distance? Gibson criticized this way of stating the problem of size constancy in its failure to acknowledge that in terrestrial environments most objects rest on *textured* ground surfaces. The same object, seen at different distances on the ground, covers (occludes) the same number of texture units regardless of distance, illustrated in Figure 1(a). Likewise, when there is a horizon line present in the field of view, it will intersect equally sized objects at the same height even when the objects are located at different distances (Figure 1b). In short, relations in the field of view (i.e., object/ground surface texture, object/horizon) serve as perceptual information for object size constancy. Because these relations remain the same under changes of viewing distance, they are invariants.

Optic Flow

As a perceiver moves forward, an optic flow of structured light from reflecting surfaces is generated, appearing to stream past the perceiver. This experience of self-motion through vision—*visual kinesthesis*—explains the commonplace illusion of feeling that you are moving when actually movement is occurring in your surround (an adjacent car in a line of traffic).

Self-generated optic flow radiates outward from a stationary point in the optic array, and this point of outflow (center of expansion) specifies where one is heading (Figure 1c). As a result, perceivers can control their direction of movement by maintaining

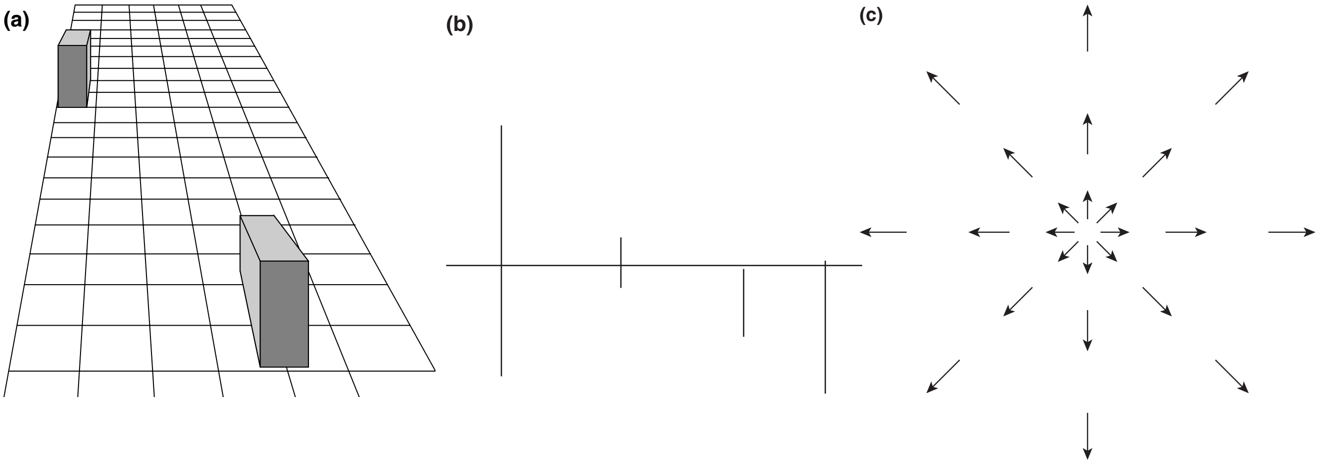


Figure 1 Perceptual Information for Depth and Size

Notes: (a) Object size—ground texture invariant: Objects of equal size occlude equal units of ground surface texture, even when they are positioned at different relative distances from the perceiver. (b) Horizon—object height invariant: The horizon line appears to intersect equally-sized objects in the same proportion, even when they are positioned at different relative distances from the perceiver. The two vertical lines on the left are depicted as twice the eye height of the observer. The two at the right are depicted as nearly the same eye height as the observer. (c) Optic flow showing the focus of expansion during approach to a flat surface.

the center of outflow on the target of interest. For example, a predator in pursuit of a prey must keep the prey at the focus of expansion of optical outflow. If the solid angle of the target changes (smaller or larger) in the optical flow field, this change is information for the predator that it is losing or gaining ground on the prey. In the case of a stationary target and a moving perceiver (or inversely a moving object relative to a stationary perceiver, as in the case of a thrown baseball), *time to contact* is specified by rate of flow. In general, time-to-contact (dubbed *tau* by David N. Lee) for a constant velocity of approach is given by the ratio of a changing quantity to its speed of change. An object in the world separated from an observer by a gap of 10 meters, traveling at 10 meters per second, will arrive in one second. This suggests that to perceive that time of arrival, an observer would need to know an object's distance and velocity as elements of a computation. However, the image of the gap on a retina, divided by the velocity of gap closure, yields the same time—without the need for separate estimates of distance and size. Lee proposed that this quantity is detected by animal nervous systems as a unit and is widely used in the control of timing.

Note that in each of these cases, perceiving is based on the *detection* of relevant higher-order information in the optic array *over time* rather than the mental calculation of separate variables, such as distance of the target and speed of movement. In this respect, self-motion and time-to-contact, like relative size constancy, is specified by information available to be *perceived* in the ambient optic array. During the past three decades, optic flow has been the subject of a great deal of research and discussion.

Occluding Edges

When a perceiver moves laterally with respect to two objects positioned one behind the other, the surface of the farther object will be gradually revealed (surface accretion) at the visible edge of the closer one. If the perceiver then reverses direction, the surface of the farther object will be concealed (surface deletion) at the occluding edge. These *reversible occlusion effects* serve as information for two perceptual phenomena: (1) the *relative distance* of two surfaces in the line of sight, and (2) the apparent *persistence* of objects that are presently not in view.

Relative Distance

A surface that is revealed or concealed at an occluding edge is experienced as being located farther away from the perceiver. In traditional theories, object occlusion is a pictorial cue for relative distance, but as a static cue (in the absence of movement), it is unreliable.

The earliest research pointing to the significance of occluding edge effects for moving perceivers was Eleanor Gibson's classic research using the *visual cliff* apparatus. Figure 2 shows a diagram of a visual cliff emphasizing the occluding edge.

A crawling baby at the edge of a visual cliff causes perceivable accretion and deletion of the lower surface at the edge of the upper surface with her movements. In this way, information for depth at an edge is available to be perceived by even newly crawling babies, and as Gibson found, they avoid venturing forth under such circumstances. This outcome contradicts what would be expected from a traditional empiricist view, which assumes that distance, not being presented in the retina, would require considerable learning. Likewise, it would be wrong to attribute this outcome to innate knowledge, as evidence with non-crawling babies has shown. Moreover, positing an innate "cognitive module," as is currently fashionable, merely begs the question of what is perceived. The individual discovers occlusion at an edge and its functional significance with exploration.

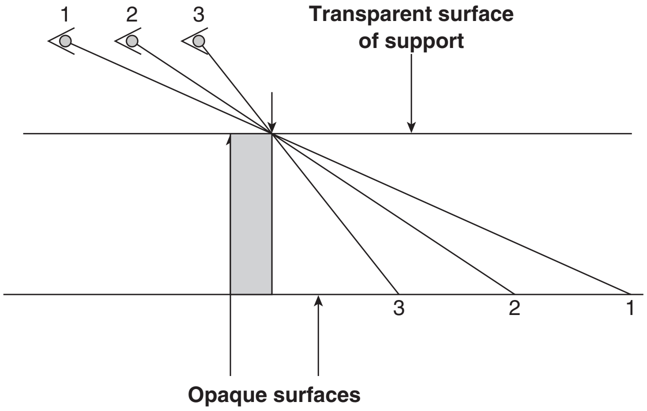


Figure 2 Occlusion Effects and the Visual Cliff

Notes: At position 1, the lower surface can be seen to the right of the numeral 1, but everything to the left is occluded. At position 2, the lower texture between 1 and 2 is revealed. At position 3, only the area to the left of the 3 is hidden. Moving from position 1 to 3, then, reveals texture (accretion) and from 3 to 1 hides texture (deletion). Thus, reversible transition of accretion or deletion of texture specifies depth at an edge.

Object Persistence

In the process of turning our heads and moving our bodies, some objects go out of sight, as others come into view. Curiously, objects that are no longer immediately visible are experienced as still being "there" (object persistence). They are not experienced as having gone out of existence, but rather as having gone out of sight. But how can this be? How is it that an object not immediately in sight can be experienced as persisting? If we bear in mind that perceiving is a process occurring over time, when surface information gradually goes out of sight at an occluding edge with perceiver movement, it can be brought back into sight by reversing the movement. This *reversible transition* at an occluding edge specifies the *persistence* of the object even when it is temporarily hidden (out of sight).

This effect has been demonstrated in a variety of ways. For example, if very young infants observe a moving object that gradually is occluded at the edge of a stationary object (it appears to go behind the latter), they will *anticipate* the reappearance of the moving object at the occluding surface's opposite edge. (Anticipation is gauged by indications of their surprise when it does *not* reappear.) From an ecological perspective, the infant's anticipation of the reappearance of an object that presently is out of sight indicates an awareness of the persistence of the hidden object. To generalize from this, it would seem that perceptual awareness of persisting objects and even places that are presently out of sight (such as the town over the next hill) is grounded in reversible occlusion effects. From this point of view, perceiving is an act of cognition.

Affordances

A number of perception researchers over the years have noted that some perceived characteristics of the environment appear to be meaningful (e.g., the cliff-edge affords falling-off, a surface at knee height appears to be a place to sit). It seems difficult, however, to reconcile apparent meaning in perception with the standard formulation of the stimulus for vision being some physical parameter of light. Light, considered as a physical stimulus, cannot carry a psychological quality such as meaning. But perhaps from the point of view of ecological optics, it can.

J. J. Gibson proposed a new concept, *affordances,* to refer to the perceived functional significance of environmental features, and he tied affordances to properties of the environment taken relative to an individual. For example, if a surface of support is roughly knee-high, it will be perceived as affording "sitting-on." Note, however, that because knee-height can vary across individuals, what might afford sitting-on for a tall person would not readily do so for a short person. This seemingly trivial fact has important implications for theory. An affordance points to a domain of properties that is not "in" the environment considered apart from an individual nor are these intra-psychic properties "in" the individual considered apart from the environment. *Affordances are relational properties.* Like optic flow and reversible occlusion, affordances arise from dynamic perceiver-environment relations.

Experimental work during recent decades has provided evidence for this supposition. For example, William Warren investigated individuals' judgments of whether a surface of support afforded stepping-up-on. He found that individuals' judgments were based on the ratio of step riser height to their leg length. Significantly, because this value is scaled relative to the body, rather than being a property of the environment independent of the perceiver, it is constant (invariant) across individuals of varying heights.

Subsequent research has examined numerous other affordances, including whether a surface affords sitting-on, whether an aperture can be passed through, whether an object is graspable as well as whether it is within reach, whether a barrier can be stepped over, and whether a task is perceived to require the participation of another person in addition to oneself. Each of these affordances is scaled relative to an individual; hence, they are not properties of the environment strictly speaking, but properties of a person-environment system.

An affordance approach to environmental properties underscores the significance of the body as a frame of reference in perception and cognition. The recent burst of activity on embodiment in cognitive and social psychology can be partially traced to J. J. Gibson's groundbreaking work on affordances, although much of this work fails to adopt the relational perspective of affordances.

Perceptual Learning and Development

The research programs of J. J. and Eleanor Gibson ran along complementary but parallel tracks. In one of their few collaborations, they proposed that growing perceptual awareness of new environmental properties over time (perceptual learning) is largely a process of discovery (perceptual differentiation) rather than of associative learning. As noted earlier, the ambient optic array is filled with light that has been structured by reflecting distal surfaces. This ecological fact indicates that a rich array of potential information is available to be perceived by individuals, and at any given moment, a perceiver is extracting only a limited portion from what is available. All of the concepts discussed in this entry refer to examples of higher-order information that are available to be perceived and that individuals become attuned to with exploration. Over time, a perceiver's awareness of properties of the environment becomes ever richer and extended with continuing opportunities to detect distinguishing relational and higher-order information specifying features, objects, and events.

Eleanor Gibson and her students have pursued these ideas through innovative developmental research, including investigations of rigid and nonrigid motion, traversability of surfaces, intermodal perception (e.g., detecting common invariants across different perceptual systems), perception of surface properties relevant to locomotion (e.g., slopes), reaching and grasping, and the use of tools. Much of this work has served as an impetus for psychology's recent wholesale reassessment of the perceptual world of infants.

Philosophical Implications: Ecological Realism

The ecological approach is unique among perceptual theories in its claim that the environment is perceived directly (*direct realism*). This view holds that our everyday impression that the environment, in most cases, is as it appears to be, is warranted. Historically, it has been pejoratively referred to as "naïve" realism. Why is that?

Traditional approaches to visual perception uniformly claim that what we experience when we perceive the environment is not the environment itself, but a mental construction of it (a view called *indirect realism*). The basis for this seemingly counter-intuitive position rests in its initial assumptions. As we have seen, traditional approaches take it for granted that the stimuli for vision are initially projected on a two-dimensional picture plane.

These stimuli are inherently equivocal (ambiguous) because as projections on a two-dimensional plane, they do *not* uniquely specify their sources in the environment. And yet it is our impression that we experience with considerable certainty a three-dimensional world around us. Hence, it must be the case—according to traditional views—that perceiving involves some intervening processes that interpret this two-dimensional display of equivocal stimuli, resulting in a "best guess" as to what the current environmental conditions truly are. The product of such processes is posited to be a mental image or a mental representation of the environment. Consequently, each individual is assumed to experience the environment as a private, intrapsychic mental construction. This assumption has created philosophical problems for centuries. Psychologists who operate from one of the traditional theories have either ignored these problems or dismissed them as being irrelevant to the research enterprise.

As we have seen, from an ecological perspective, perceiving is a dynamic process of detecting information in the stimulus array that univocally specifies features of the environmental layout. To review, the structured light of the ambient optic array carries information specific to environmental objects and events. Critically, actions of the individual introduce perceivable changes in the ambient optic array, thereby revealing non-change (invariant structure) in the context of change. Because these higher-order invariants are specific to particular stable and persisting features of the environment, they function as information for direct perception of those features. Reciprocally, change generated by the perceiver's actions specifies the self as a source of agency (intention).

This is not to say the perceiving is always free of error—under special circumstances the perceivable structure in the light can mislead, as in the case of an apparently bent stick in water. Nor is this to deny that sometimes individuals are forced to engage in guesswork when there is insufficient or equivocal information (e.g., conditions of fog, inadequate illumination), or under conditions of artificial constraints imposed in research laboratories. Moreover, all of us rely to a great extent on second-hand descriptions (e.g., reportage) or representations (e.g., photographs) of the actual state of things. This second-hand information is accurate only to the degree that it faithfully carries information about its sources. Information in the ambient array affords us the opportunity to perceive the environment directly and to look for ourselves.

William Mace and Harry Heft

See also Amodal Perception; Causality; Direct Perception; Event Perception; Navigation Through Spatial Layout

Further Readings

Flach, J., Hancock, P., Caird, J., & Vicente, K. (Eds.). (1995). *Global perspectives on the ecology of human-machine systems* (Vol. 1.). Hillsdale, NJ: Lawrence Erlbaum.

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development.* New York: Oxford University Press.

Gibson, J. J. (1950). *The perception of the visual world.* Boston: Houghton Mifflin.

Gibson, J. J. (1966). The se*nses considered as perceptual systems.* Boston: Houghton Mifflin.

Gibson, J. J. (1986). *The ecological approach to visual perception.* Mahwah, NJ: Lawrence Erlbaum. (Original work published 1979)

Hecht, H., & Savelsbergh, G. J. P. (Eds.). (2004). *Time to contact.* Amsterdam: Elsevier.

Heft, H. (2001). *Ecological psychology in context.* Mahwah, NJ: Lawrence Erlbaum.

Pepping, G.-J., & Grealy, M. A. (Eds.). (2007). *Closing the gap: The scientific writings of David N. Lee.* Mahwah, NJ: Lawrence Erlbaum.

Reed, E. S. (1996). *The necessity of experience.* New Haven, CT: Yale University Press.

Warren, W. H., Jr. (Ed.). (1998). Visually controlled locomotion and orientation. *Ecological Psychology, 10,* 157–346.

EFFORT: PERCEPTION OF

Perception of effort, also known as perceived exertion or sense of effort, refers to the conscious sensation of how hard, heavy, and strenuous a physical task is. This perception depends mainly on feelings of effort in the active limbs, and the sensation of heavy breathing (a type of dyspnea). Several authors think that somatic sensations such as muscle pain, thermal discomfort, and chest pain also contribute to perceived exertion. However, it has been demonstrated that people can differentiate perception of effort from these somatic sensations experienced during exercise (e.g., muscle pain). This entry describes the history, measurement methods, physiological and psychosocial determinants, neurophysiology, effort and other perceptions, and relevance and practical applications of perception of effort.

History and Measurement Methods

The sense of effort, including perception of force and weight, has been studied by psychologists and physiologists since the mid-to-late 19th century. However, the construct of perceived exertion was introduced in the early 1960s by Gunnar Borg, a Swedish psychophysicist with an interest in exercise. Borg also developed the two most common instruments used to measure perception of effort: the rating of perceived exertion (RPE) scale and the category-ratio (CR-10) scale. Both scales are based on the principle that most people share meanings of "no exertion" and "maximal exertion" based on similar experiences with physical activity, and careful selection of verbal expressions of effort to obtain category scales with interval or ratio properties.

The RPE scale measures the intensity of perceived exertion on a 15-point category scale representing equal intervals and ranging from 6 ("no exertion at all") to 20 ("maximal exertion"). Seven of the 13 intermediate points are anchored to verbal expressions of effort such as "very light," "somewhat hard," and "heavy." The CR-10 scale can be used to rate not only perception of effort, but also the intensity of other sensations such as pain. This scale ranges from 0 (no perception) to 10 (the strongest perception ever experienced) with various verbal anchors (e.g., "weak" and "moderate"). Subjects can give ratings higher than 10 (e.g., 13) if they perceive an intensity stronger than the one they have ever experienced. With proper familiarization and standardized instructions, both instruments have good reliability and validity. The latter has been demonstrated by significant correlations between RPE and objective measures of physiological strain (e.g., heart rate and blood lactate concentration) during exercise with incremental workloads.

Although the RPE scale and the CR-10 scale are the most widely used, other instruments have been developed to rate perception of effort. These methods include visual analog scales and pictorial scales portraying people exercising at different intensity levels. Classical psychophysical methods such as magnitude estimation, weight discrimination, and force-matching have also been used to quantify perceived exertion.

Physiological and Psychosocial Determinants

The workload at which somebody is exercising (e.g., running speed, resistance on the stationary bike, or amount of weight lifted) is an important factor determining perception of effort. Indeed, perceived exertion increases as workload increases. However, across individuals, there is a poor relationship between absolute workload and perception of effort. This is because perceived exertion accurately reflects relative exercise intensity, which depends on two factors: absolute workload and individual exercise capacity (i.e., physical fitness). For example, the same running speed of 12 kilometers per hour (km/h) can be perceived as "very hard" by an unfit middle-age man or as "very light" by an endurance athlete. However, both subjects would perceive a similar effort when running at 80% of their maximum endurance running capacity. Because physical fitness is affected primarily by physical training and health status, these two factors are important determinants of perceived exertion. Another important determinant of perceived exertion is the duration of the exercise bout. Perception of effort is known to increase significantly over time during prolonged exercise at a fixed workload. This phenomenon is an essential feature of fatigue during physical tasks. Nutritional (e.g., caffeine and muscle glycogen) and environmental factors (e.g., altitude and ambient temperature) are among many other physiological factors affecting perceived exertion.

Like other subjective feelings, perception of effort or its rating can be affected by psychological factors. These include personality, mood, somatic perception, locus of control, and self-efficacy. The presence and gender of another person during exercise testing can also influence RPE, but the effects of other social factors are poorly understood at present. Both psychological and social factors seem to affect perception of effort at low-to-moderate exercise intensities rather than during high-intensity exercise.

Neurophysiology

The neurophysiology of perceived exertion has not been extensively investigated, and its exact mechanisms are still being debated. One point of view is that perception of effort results from the complex integration of different afferent sensory inputs to the brain including those underlying proprioception, pain, and thermal discomfort. This proposal is supported by the significant correlations between RPE and several indicators of physiological strain measured during exercise (e.g., heart rate and blood lactate concentration). However, these high correlations may not be indicative of a cause-and-effect relationship. Indeed, because RPE increases as workload increases, any physiological variable that increases with workload could be related to RPE. Importantly, perception of effort is not affected by interventions (e.g., epidural anesthesia and lung transplantation) that block afferent feedback from receptors sensing the physiological conditions of skeletal muscles and internal organs.

An alternative explanation is that perception of effort reflects the magnitude of the central motor commands sent from the brain to the active muscles, including the respiratory muscles. This is achieved by forwarding neural signals, termed *corollary discharges,* from motor to sensory areas of the cerebral cortex. This view is supported by several experimental studies in which peripheral neuromuscular function is reduced using, for example, partial curarization or muscle fatigue. In these experimental conditions, higher perception of effort is thought to reflect the increased central neural drive necessary to exercise at the same workload with weaker muscles.

Experimental studies using hypnotic suggestion, dissociative cognitive strategies, imagery, psychoactive drugs, and mental fatigue suggest that cognitive factors are also important for perception of effort. Some of these experimental manipulations have also been used to locate the cortical areas associated with RPE using positron emission tomography. These areas include the anterior cingulate cortex, insular cortex, thalamus, and supplementary motor area.

Effort and Other Perceptions

Perception of effort interacts with other sensory information (e.g., proprioception and optical information) to produce other perceptions. For example, experimentally induced muscle weakness influences the perception of static limb position because more effort is required to hold the limb against gravity. When the limb is not held against gravity (no effort required), muscle weakness does not influence judgments of static limb position.

Experimental manipulations of effort also affect perception of distance. For example, wearing a heavy backpack significantly increases subjective judgments of distance to various targets. Interestingly, in this and other experiments, higher effort was anticipated. Therefore, effort does not need to be exerted to affect perception of distance.

Relevance and Practical Applications

Perceived exertion is a common phenomenon in daily life (e.g., climbing a flight of stairs or lifting a heavy object). Therefore, this perception plays an important role in regulating our physical activity behaviors. These behaviors range from choosing the pace during endurance competitions to adopting a sedentary lifestyle. It has also been proposed that primary disturbances in the sense of effort may underlie some of the symptoms characteristic of chronic fatigue syndrome and schizophrenia.

From an applied perspective, RPE is widely used as an adjunct to physiological and clinical measures (e.g., electrocardiogram) during maximal exercise tests. It can also be used to predict exercise capacity from responses to submaximal exercise tests. Because it accurately reflects relative exercise intensity, perception of effort is useful to prescribe and monitor exercise intensity in individuals participating in endurance and team sports, cardiac rehabilitation programs, and fitness training. The RPE scale has also been applied to evaluate work demands in a variety of occupations involving significant physical effort.

Samuele Marcora

See also Brain Imaging; Corollary Discharge; Kinesthesia; Pain: Physiological Mechanisms; Proprioception; Psychophysical Approach; Scaling of Sensory Magnitude; Temperature Perception

Further Readings

Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Medicine & Science in Sports & Exercise, 14,* 377–381.

Enoka, R. M., & Stuart, D. G. (1992). Neurobiology of muscle fatigue. *Journal of Applied Physiology, 72,* 1631–1648.

Hamilton, A. L., Killian, K. J., Summers, E., & Jones. N. L. (1995). Muscle strength, symptom intensity, and exercise capacity in patients with cardiorespiratory disorders. *American Journal of Respiratory and Critical Care Medicine, 152,* 2021–2031.

Marcora, S. M. (2009). Perception of effort during exercise is independent of afferent feedback from skeletal muscles, heart and lungs. *Journal of Applied Physiology.* DOI: 10.1152/japplphysi01.90378.2008

Morgan, W. P. (1994). Psychological components of effort sense. *Medicine & Science in Sports & Exercise, 26,* 1071–1077.

Noble, B. J., & Robertson, R. J. (1996). *Perceived exertion.* Champaign, IL: Human Kinetics.

Proffitt, D. R., Stefanucci, J., Banton, T., & Epstein, W. (2003). The role of effort in perceiving distance. *Psychological Science, 14,* 106–112.

ELECTRONIC NOSE

Across a wide range of applications there is a need to rapidly detect, identify, and quantify complex volatile mixtures. Food manufacturers need to monitor the quality of incoming raw products and evaluate the perceptual properties of their products; environmental agencies seek to regulate the emission of malodors from waste treatment and other industrial plants. Humans readily accomplish these tasks through the sense of smell. Could an instrument be developed to mimic these capabilities? This entry describes the technology of electronic noses, discusses the challenges associated with mimicking human olfaction, and reviews past and current application areas of electronic noses.

Definition

Electronic noses (or e-noses for short) are scientific instruments that detect, recognize, and quantify volatile chemicals in a manner that is loosely reminiscent of the sense of smell. The concept of an e-nose dates back to a 1982 article by Krishna Persaud and George Dodd, which showed that fine discrimination between odorants is possible without highly selective sensors. Instead, and much like the human nose, e-noses rely on cross-selective sensors; that is, sensors that respond to many different odorants, though the strength of the response varies with each odorant. To discriminate odorants, e-noses (and human noses as well) use an array of sensors with different cross-selectivities. In this way, the response pattern across sensors becomes a fingerprint for each particular odorant. In addition to chemical sensors, e-noses require a pattern-recognition system and an odor-delivery system.

Chemical sensors consist of two elements: a sensing layer and a transducer. The sensing layer is made of a material whose properties change when odorant molecules are adsorbed on its surface; as an example, the sensing material may change its electrical resistance or its mass. In turn, the transducer converts these changes into an output signal; as an example, a transducer may convert mass changes in the sensing layer into frequency changes of an output signal, which then can be easily recorded with a computer. Sensing materials for e-noses can be inorganic (metal oxides are typical), organic (e.g., conducting polymers), and biological (e.g., proteins). Typical transducers for e-noses are based on conductivity or mass changes, though other transduction principles can be used as well (e.g., capacitive, calorimetric, optical, field effect), sometimes simultaneously.

Most e-noses operate under two basic forms of odor delivery: static headspace and dynamic headspace. In static headspace analysis, a sample is placed inside a closed container along with the sensor array, and the headspace above the sample is allowed to equilibrate before recording the sensor array response. In the dynamic headspace method, an effluent (e.g., filtered room air) flows through the sample continuously, carrying odorant molecules to the sensor array located in a separate chamber. Several techniques (e.g., preconcentration, separation) can also be used to condition odorant samples before they are delivered to the sensors.

Once an odorant sample has been delivered to the sensor array and the corresponding response has been recorded, the signals are analyzed with a pattern-recognition system. Depending on the application, the goal of the analysis may be to classify the sample into one of several potential categories (e.g., fresh vs. spoiled milk), predict a set of properties for the sample (e.g., concentration, quality), or find samples that are similar (e.g., clustering).

Mimicking Human Olfaction

The term *electronic nose* suggests that the instrument could be used to predict the perceptual properties of an odorant. Prediction of perceived intensity is straightforward because both chemical sensors and olfactory receptors have monotonic concentration-response curves. However, prediction of qualitative descriptors is more challenging, unless the problem is constrained to specific applications (e.g., rating the freshness of particular foods). Unlike vision and hearing, in which the properties of rods/cones and hair cells can be explained in terms of light wavelengths and sound frequencies, respectively, the primary dimensions of olfaction are still unknown. Findings by Linda Buck and Richard Axel, which led to the 2004 Nobel Prize in Physiology or Medicine, indicate that olfaction has a large number of dimensions (about 350 olfactory receptors in humans). However, the precise odorant properties detected by each olfactory receptor are yet unknown, which makes it difficult to develop chemical sensors with response properties similar to those of olfactory receptors. In addition, olfactory stimuli can contain many individual chemicals (e.g., more than 800 different volatile compounds in coffee), and the perception of odor mixtures is poorly understood. Finally, current chemical sensors have detection thresholds that are several orders of magnitude higher than those in the olfactory system.

Applications and Outlook

E-noses represent a trade-off between high-end analytical instruments, such as mass spectrometers or gas chromatographs, and sensory analysis with human panels. Like analytical instruments, e-noses can provide objective results; like human panels, e-noses provide an overall response (a gestalt), rather than a fine analysis of a smell into its individual constituents. Early commercial e-noses were marketed as general-purpose instruments that could be calibrated to any application by “retraining” their pattern-recognition model on the odorants of interest. Experience has shown that this is not possible with current chemical sensors. As a result, current commercial efforts are moving toward application-specific chemical sensor systems. In the food and beverage industry, e-noses can be used to inspect raw ingredients (e.g., olive oil quality), monitor production processes (e.g., fermentation and roasting), and estimate the shelf life of products (e.g., fish spoilage), to name a few scenarios. E-noses also have application in environmental problems, where they could be used to monitor emissions (e.g., waste disposal sites) and indoor air quality (e.g., ventilation systems in hospitals). Finally, e-noses have shown potential for the rapid diagnosis of disease by analyzing odors from skin (e.g., yellow fever), breath (e.g., diabetes), sweat (e.g., diphtheria), as well as odors from waste products and excretions. Thus, although e-nose systems are yet unable to describe a Marqués de Murrieta 1989 Ygay Rioja wine as having “intense black fruit and vanilla aromas with a lovely scent of something surprising like toasted coconut,” a variety of applications can benefit from specialized e-noses that can rapidly recognize and quantify volatile compounds.

Ricardo Gutierrez-Osuna

See also Ageing and Chemical Senses; Air Quality; Olfactometry; Olfactory Quality; Olfactory Receptors and Transduction; Olfactory Stimulus

Further Readings

Gutierrez-Osuna, R. (2002). Pattern analysis for machine olfaction: A review. *IEEE Sensors Journal, 2*(3), 189–202.

Nagle, H. T., Schiffman, S. S., & Gutierrez-Osuna, R. (1998). The how and why of electronic noses. *IEEE Spectrum, 35*(9), 22–34.

Rock, F., Barsan, N., & Weimar, U. (2008). Electronic nose: Current status and future trends. *Chemical Reviews, 108,* 705–725.

Sneath, R. W., & Persaud, K. (2003). Correlating electronic nose and sensory panel data. In T. C. Pearce, S. S. Schiffman, H. T. Nagle, & J. W. Gardner (Eds.), *Handbook of machine olfaction* (pp. 377–397). Weinheim, Germany: Wiley-VCH.

ELECTRORECEPTION

Electroreception is the sensory ability of some species of animals to detect weak, naturally occurring electrostatic fields in their environment. It is found in the most ancient lineages of fishes and survives today in about 10% of all vertebrate species—all of which are aquatic. Electroreception was lost in evolution when vertebrates crawled onto land, so we humans can only imagine the perception of an electrical world with this sixth sense. For reasons unknown to biologists, electroreception was also lost among the ancestors of the modern fishes only to reappear in two separate lineages of ray-finned, or teleost fishes. The sense also appeared independently with the evolution of the monotreme mammals such as the platypus and echidna. All electroreceptive species possess sensitive electroreceptor organs in their skin that can detect electric fields as weak as a fraction of a microvolt per centimeter. These electroreceptors exhibit a low-resistance pathway for current to flow from the external environment to a sensory epithelium containing cells with voltage-sensitive ion channels in their membranes. As these sensory cells are depolarized, they activate nerve fibers to send action potentials to the brain. The oldest lineages of fishes with electroreceptors have ampullary organs that can sense direct current (DC) fields. This allows them to detect and capture prey in the dark or in murky water. The more recently evolved teleosts with electroreceptors have tuberous organs that can sense alternating current (AC) fields, an ability that is useful for social communication and for active object-location if it is combined with discharges from an electric organ. Terrestrial animals have no use for electroreception because the high resistance of the air prevents the flow of electric current.

Only a few investigations have searched for electroreception among invertebrates, although a few studies report on electric sensitivity of crayfish and worms, but the behavioral thresholds are so high (more than 3 volts per centimeter (V/cm) in the worm *Caenorhabditis elegans*) that it is doubtful that natural fields would ever activate them. For comparison, sharks and rays, which have the most sensitive ampullary receptors, have thresholds as low as 0.02 microvolts per centimeter, the equivalent of an electric field set up between one pole of a 1.5 volt battery in San Francisco Bay and the other pole connected to the ocean near Los Angeles.

Even though humans lack electroreceptors, we can perceive strong electric currents from batteries, static generators, and other sources if we make direct contact with them or establish indirect contact through a conducting medium such as water. The perception is unpleasant because electric currents stimulate sensory and motor nerve fibers indiscriminately, producing a tingling sensation at threshold levels, and painful sensations and muscle twitch at higher levels. Human thresholds are in the 1 milliamp current range, which translates into electric fields of 0.1 to 1 Volts/cm in freshwater. Our threshold for sensing an electric field in air is more than 4 kilovolts per centimeter (kV/cm), a perception that originates with the mechanical displacement of charged dermal hairs in electric fields. This entry describes the discovery, perception, and functions of electroreception, as well as the evolution and anatomy of electroreceptors.

Discovery

Electroreceptor organs were first identified physiologically in the early 1960s from weakly electric fish by the U.S. neuroscientist Theodore H. Bullock and colleagues and by French scientists Thomas Szabo and A. Fessard. The existence of electroreceptors had been anticipated in the late 1950s by Hans W. Lissmann from Cambridge University who was the first to discover continuous “weak” electric discharges from an electric organ—the tail of the African fish, *Gymnarchus niloticus*. By 1958, he had demonstrated the reason for the discharge by showing that these fish could detect the presence of glass and metal rods or other conducting or nonconducting objects at distances of 10 centimeters or more, even in the absence of visual, mechanical, or chemical cues. Lissmann postulated that they were sensing the distortions of their own electric organ discharges as electrical shadows on their skin. Lissmann correctly guessed that there were dermal electroreceptors in these fish. He called their behavior “electrolocation” after the well-known ability of bats to echolocate objects by detecting returning ultrasonic echoes from their calls.

In the 1960s, Dutch scientists Sven Dijkgraaf and Adranius Kalmijn established that sharks and rays, which have dermal sense organs called ampullae of Lorenzini, could sense weak electric currents from their prey organisms such as flatfishes even when they were buried under sand. Dijkgraaf and Kalmijn showed that the ampullae of Lorenzini were essential to this behavior, which was entirely based on electrosensory cues, and that prey have weak DC electric fields surrounding their gills, gut, and skin wounds that gave away their presence to the sharks. When given a choice between a prey fish covered with plastic wrap and a pair of wire electrodes connected to a prey fish in another tank, the shark preferred to dig up the wires. Kalmijn and colleagues called this ability "passive electrolocation" in contrast to the "active electrolocation" ability of *Gymnarchus niloticus* discovered by Lissmann.

Evolution and Anatomy of Electroreceptors

Electroreception first appears in crainates among lampreys that have epidermal "end-bud" organs innervated by the lateral line nerve. The end-bud receptors are sensitive to weak, low-frequency (DC to 50 hertz [Hz]) electric fields. Among the earliest jawed vertebrates we see the first ampullary electroreceptors embedded in the skin or at the base of a long conducting canal leading from the skin surface to a specialized patch of modified sensory epithelial cells in the case of ampullae of Lorenzini. The sensory epithelial cells resemble hair cells of the lateral line or inner ear, and the nerve fibers connecting them to the brain travel in the same nerve bundles with the lateral line nerves. The cells lining the ampullary canals are packed tightly to make a high resistance insulator around the low-resistance canal lumen, which is composed of a salty gelatinous matrix. Like an insulated wire, this arrangement is ideal for conducting electric currents to the sensory cell membranes. All vertebrate electroreceptors follow this basic design.

Electroreception is shared by all of the primitive aquatic vertebrates, including some aquatic amphibians, but it was lost in the amniotes as they made the transition to a terrestrial existence. In those monotreme mammals in which the sense reappears, electroreception is derived from modified mucous glands on the bill or snout. These mucous glands have low-resistance canals leading from water to bare nerve endings of the trigeminal nerve. Electroreception is absent in most modern fishes (including bowfins, gars, and teleost fishes) except that it appears in two independently evolved lineages of teleosts, first as ampullary receptors among catfishes and independently among notopterid knifefishes of Africa. Remarkably, tuberous or AC sensitive electroreceptors also appear in both of these lineages as subgroups of electric fishes. These are the South American gymnotiform fishes, related to the electric eel, and the unrelated mormyroid or elephant-snouted fishes from Africa. Members of both groups use their tuberous organs for active electrolocation of objects and for electrical communication. There are sufficient anatomical and physiological differences between teleost electroreceptors and the ancient ampullae of Lorenzini so that biologists conclude they were derived secondarily. The molecular basis for electrosensory transduction is unknown for electroreceptors.

Functions of Electroreception

Ampullary electroreceptors are used for passive electrolocation of the DC electric fields from prey organisms, predators, or conspecifics. The ability to sense other animals gave these fishes a decided advantage in navigating in the dark or in turbid water. Tuberous electroreceptors, which are found only in the two independent lineages of teleosts already mentioned, are used for sensing the AC electric fields from electric organs, either in the context of active electrolocation of objects, or for social communication. Electric communication signals from electric organs can be quite elaborate and varied. A rich repertoire of electric signals is used for signaling threats, submission, alarm, and courtship, and some of these fish have electric signals for sex, species, and individual recognition.

Perception

Because humans lack electroreception and the words to describe it, we can only imagine what this sense must be like for other animals. For those with true electroreceptors, the sensation is probably nothing like the unpleasant feeling we associate with electric shocks. To understand what an animal perceives, we can only do behavioral tests, make physiological recordings, and examine behavior of animals with sensory deficits, and even so we will never experience what the animal feels. We imagine that electroreception must resemble the touch because both modalities have receptors distributed over the body surface and both have sensory maps of the body surface in the brain. However, electroreception is also like the thermal sense because electric currents flow in vague and diffuse pathways and never create sharp boundaries. But there are also similarities between electroreception and hearing because both modalities are encoded by frequency-tuned receptor cells resembling hair cells and both use time encoding to represent time varying stimulus waveforms. Electroreception is also similar to vision because multiple receptor types are distributed over the skin surface as in the retina, and in both modalities, there are local regions where receptors are extremely dense, creating a fovea of acute spatial resolution. Whether electroreception is perceived more like we perceive touch or hot or cold or hearing or seeing is unimportant to the scientists who strive to describe the sensory experience in terms of physical attributes of the stimulus, electrophysiological properties of sensory afferents, or central neurons, or the behavioral performance during natural behavior and trained discrimination. An animal's perception will remain unknown to us. Nevertheless, the widespread appearance of electroreception among aquatic vertebrates and the multiple cases of its evolution suggest that this sense, although foreign, is an extremely important part of the evolutionary history of our vertebrate lineage.

Carl D. Hopkins

See also Cutaneous Perception: Physiology; Echolocation; Evolutionary Approach; Evolutionary Approach: Perceptual Adaptations

Further Readings

Albert, J. S., & Crampton, W. G. R. (2005). Electroreception and electrogenesis. In D. H. Evans & J. B. Claiborne (Eds.), *The physiology of fishes* (pp. 429–470). New York: CRC Press.

Bullock, T. H., Hopkins, C. D., Popper, A. N., & Fay, R. (Eds.). (2005). *Springer handbook of auditory research: Electroreception.* (Vol. 21). New York: Springer.

Moller, P. (1995). *Electric fishes: History and behavior.* New York: Chapman & Hall.

EMBODIED PERCEPTION

Embodied perception refers to a conglomerate of theoretical claims postulating that the body, its movements, and the interaction with the environment fundamentally shape people's perception of the world. All these claims have in common that they criticize "disembodied" notions of perception that are characteristic of the information-processing approach. Conceived during the cognitive revolution that made the computer the main model for the mind, the information-processing approach postulates that the function of perception is to encode and classify physical stimuli to form mental representations. These representations are then used in further cognitive (symbolic) processing. However, different strands of embodied perception take issue with the information-processing approach for different reasons. As yet, there is no set of core assumptions that every embodied perception researcher would subscribe to. This entry describes four influential notions of embodied perception by classifying them in terms of their core assumptions.

Focus on Interaction With the Environment

An influential approach to embodied perception has its roots in ecological psychology and the work of J. J. and Eleanor Gibson. At the core of (neo)-Gibsonian versions of embodied perception is the idea that observer and environment are so tightly coupled that a meaningful explanation of perception must focus on their interaction. Perception does not happen in the mind or brain alone. Accordingly, this approach rejects the assumption of the information-processing approach that unconscious inferences create a stable mental representation of the stimuli that appear in the outside world.

Instead, it is postulated that to understand perception one needs to investigate how the relation between observer and environment enables *direct perception.* This term refers to the claim that information emerges in a dynamic spatiotemporal relationship between observer and environment, characterized by perceptual invariants—properties of the environment that remain constant as the observer moves. For example, an observer can directly perceive how large an object is from the dynamic pattern of information that the movement generates as the observer is approaching the object (optic flow). Similarly, if one is passively moved (e.g., riding a train), one can directly perceive the distance of different objects (e.g., trees along the tracks and distant hills) based on how quickly they pass by.

The ecological approach is radically embodied because it implies that perception cannot exist independently of an observer who moves through the environment. The key concept of affordance provides a further illustration of this idea: Affordances are conceived of as action possibilities that objects in the environment provide for the moving observer. Importantly, the relation between an actor's capabilities and physical properties of objects determines what is afforded, rather than the objects alone. For example, for adults, a chair affords sitting, whereas for an infant it may afford drawing himself or herself up to standing position.

The notion of direct perception has had a major impact on new developments in artificial intelligence, and especially, robotics. Earlier attempts at building intelligent machines relied on the assumption that a central processor stores input and performs logical operations that lead to a meaningful output. Embodied, interactionist thinking in robotics led to the construction of robotic agents that avoid central planners. Instead, they rely on direct links between input (perception) and actions. These links define which action is appropriate given a particular input. Conflicts between different possible actions are not solved through explicit selection of a behavior by a central controller. Rather, they are solved implicitly by organizing the different perception-action links in a hierarchical manner so that certain behaviors are prioritized over others. Despite the lack of a central controller, robots with such a subsumption architecture evolve complex, intelligent behavior in real-time through interaction with a rich environment. An illustration of this is Herbert, an MIT robot from the 1980s. Herbert was able to collect empty cans without stored long-term plans or a model of the world, relying just on a small set of behaviors such as object avoidance and locomotion.

The direct perception approach to embodiment shares with other embodiment theories a focus on the real-time character of perception, and on the emergence of adaptive behavior through close interaction with the environment. However, the outright rejection of mental representations and their neural correlates is a contentious issue. For instance, the philosopher Andy Clark, one of the main proponents of embodied perception, has argued that there must be room for at least action-centered representations and partial models of the world. The rejection of any central, integrated processing is incompatible with evidence that the brain integrates multiple sources of information and ignores the existence of internal models that provide links between perception and action in real time.

Focus on Evolution and Body Morphology

Arguments for another biologically inspired take on embodied perception come from evolutionary approaches stressing the fact that brain and body have evolved together over time to maximize the chances for survival in particular environmental niches. This fact, it is argued, has been neglected by the information-processing approach to perception where evolutionary considerations and evolved anatomical constraints have only played a marginal role.

Evolutionary approaches to embodied perception often focus on the relation between body morphology and perceptual abilities in different species. It is thought that ecology can provide important clues to how evolutionary constraints have shaped sensory systems. For instance, predators' eyes tend to be closer together and more to the front to provide the accurate binocular vision that is important for locating prey. In contrast, animals that are more likely to be preyed on have the eyes farther apart and more on the sides to produce the wide field of vision needed to spot predators.

In addition to considering the relation between body morphology and the senses, evolutionarily inspired embodiment approaches also address the relation between body morphology and neural control. For example, if a ferret's brain is rewired so that retinal projections reach the auditory cortex, the auditory cortex contains a map of visual space and its cells have properties typical of cells in a normal primary visual cortex. This suggests that the patterns of maps in sensory cortices are not hardwired, but emerge as a consequence of the patterning inherent in the sensory input (are plastic).

Evolutionary takes on embodied perception have an increasing impact on robotics and have led to the development of "evolutionary robotics." Whereas early work in this area focused on central control processes and ignored body morphology, recent studies have addressed how evolving mechanical "bodies" constrain behavior and vice versa. Using known genetic principles of reproduction, computer simulations of evolutionary processes across multiple generations have shown that simple organic units group to form stable bodies. These exhibit rather complex behaviors in response to particular environmental conditions, for example, pushing blocks out of the way. In such scenarios, genes do not explicitly code for the body structure of the evolving agents. Rather, the interaction between genes and environment determines which anatomical structures and which behaviors provide the best chances of survival.

The evolutionary take on embodied perception is close in spirit to interactionist strands of embodiment described earlier because the interaction between organism and environment is regarded as crucial. However, evolutionary approaches consider much larger time scales than are considered in ecological psychology, thus adding a historical perspective on emergent properties that arise through the interaction between organisms and the environment.

Focus on Proprioception

The label of *embodiment* is also increasingly found in more traditional perception research that is not committed to strong interactionist or evolutionary notions of embodiment. Here, embodiment reflects the view that the body is a special object of perception. Accordingly, proponents of this view maintain that perception critically depends on sensory information about one's body, especially proprioception.

The most radical version of this claim is that proprioceptive information about body movements and touch provides an inner perspective through which all other sensory information, especially visual and auditory information, appears to the observer. This inner body perspective is thought to be at the core of all perceptual experiences. Less radical versions focus less on this experiential aspect and more on specific empirical questions. Both have in common that they argue that the body senses have been undervalued in perception research.

Closely related is a renewed interest in multimodal integration. Research in this field suggests that the body can constrain perception via associations between body senses and distal senses such as vision and audition. Examples for such links include differences in perceptual ability in the area immediately surrounding one's body (near space) compared with the area outside of one's reaching space (far space), and increased perceptual efficiency in areas that are close to one's body parts. A related discussion addresses whether there are specialized processes and brain systems for the visual processing of the human body structure. For example, the extrastriate body area (EBA) that contains neurons responding to bodies may selectively filter out body-like visual configurations and pass them on to higher brain systems. Alternatively, people may map others' visually perceived bodies onto their own multimodal body schema assumed to reside in the inferior parietal cortex.

Focus on Intentional Action

Another take on embodiment comes from accounts of perception capitalizing on intentional action and the action repertoire of the perceiver. These accounts go back to William James's ideomotor theory that explains voluntary action as driven by perceptual images. Researchers criticize the neglect of intentional action in perception research. In line with interactionist notions of embodiment, perception is conceived of as serving action rather than providing an image of the world. However, the notion of internal representation is not rejected. Rather, internal representations are assumed to play a central role in controlling actions.

The basic functional claim of this approach is that actions are coded by their perceptual consequences. This creates a common representational space for performed and perceived actions (common coding). For instance, when a person claps his or her hands, this will produce a particular sound. When another person perceives this sound, this will activate a representation of the corresponding clapping action. The more similar the sounds that both persons produce when clapping, the higher this activation will be.

Support for the assumption of such perception-action links has been provided by the finding of mirror neurons in monkeys and an analogous mirror system in the human brain. The same motor areas of the brain are activated when individuals perform an action and when they perceive somebody perform a similar action. The more one is an expert at performing a particular action, the more responsive motor areas in the brain are to perceiving this action. For instance, primary motor areas are active in pianists when they hear piano tones but not in novices. Both the functional principle of common coding and the neural evidence for a mirror system define a brand of embodiment that stresses that the ability to perform intentional actions fundamentally shapes people's perception of each other's actions.

Natalie Sebanz and Guenther Knoblich

See also Action and Vision; Body Perception; Direct Perception; Mind and Body; Motion Perception; Multimodal Interactions: Visual–Haptic; Navigation Through Spatial Layout

Further Readings

Churchland, P. S. (1986). *Neurophilosophy: Toward a unified science of the mind brain.* Cambridge: MIT Press.

Clark, A. (1997). *Being there: Putting brain, body and world together again.* Cambridge: MIT Press.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Klatzky, R. L., Behrmann, M., & MacWhinney, B. (Eds.). (2008). *Embodiment, ego-space, and action.* Mahwah, NJ: Lawrence Erlbaum.

Knoblich, G., Thornton, I., Grosjean, M., & Shiffrar, M. (Eds.). (2006). *Perception of the human body.* New York: Oxford University Press.

O'Reagan, J. K., & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24,* 939–1011.

Pfeifer, R., & Bongard, J. C. (2006). *How the body shapes the way we think: A new view of intelligence.* Cambridge: MIT Press.

Schuetz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: Action-induced modulation of perception. *Trends in Cognitive Sciences, 11,* 349–355.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review, 9,* 625–636.

EMOTIONAL INFLUENCES ON PERCEPTION

The effect of emotion on perception has been a topic of interest through much of the history of psychological thought. The notion that what we see is influenced by our internal emotional and motivational states is appealing to those inclined to view aspects of the mind as interconnected even at the earliest information-processing stages. Although accumulated evidence demonstrates that emotion affects the result of perceptual processing—that is, perceptual awareness—questions remain debated and unresolved regarding the precise processing stages at which emotion exerts its influence.

In the 1940s and 1950s, pioneering work by Jerome Bruner and colleagues inspired what became known as the "New Look" movement, a loosely knit effort among psychologists to reveal contributions of emotion, knowledge, personality, and motivation to perception. In one of their classic studies, children were asked to adjust the size of a patch of light so that it matched the size of either various nearby coins or size-matched cardboard disks. The children's errors were greater when estimating the sizes of coins than cardboard disks, a pattern particularly evident among the poorer than more wealthy children. The researchers concluded that perception is influenced by the value accorded to aspects of the environment, and they made the rallying argument that motivational factors need to be considered to understand perception in the real world. Two ideas that emerged from this movement were those of *perceptual defense* and *perceptual vigilance,* which referred to the peculiar manner in which people appeared to exhibit, respectively, impaired or enhanced perception of taboo or emotional stimuli (for example, worse or better recognition of such stimuli, relative to neutral stimuli, in noisy displays). Notably, such ideas—proposed before the maturation of information-processing approaches to perception, which involved differentiation and identification of interacting perceptual stages—seemed to contain within them an insurmountable paradox: How could emotional stimuli gain distinctive perceptual status when the act of prioritizing them necessitated that their emotional significance had already been perceived?

Modern research has at least partly resolved this dilemma, demonstrating that perceptual awareness reflects the output of many processing stages, with contributions from neural regions distributed throughout the brain. Thus, there are many opportunities for the emotional significance of sensory information to be registered before awareness. The amygdala, a neural structure heavily implicated in emotional processing, appears to receive some sensory inputs that bypass cortical areas associated with attention and awareness—which may explain findings that it responds even to emotional stimuli that people can't report. Furthermore, while this structure receives input from multiple brain areas, it also projects back reciprocally even to early neural processing regions—in the case of vision, as far back as the primary visual cortex. Such feedback connections may be important mechanisms underlying the enhanced activity, as revealed through neuroimaging, in vision-related brain areas in response to emotional, relative to non-emotional, stimuli. For example, emotionally expressive faces have been found to elicit more vigorous activity in a face-selective region of the temporal lobe (i.e., the fusiform face area, or FFA) than do neutral faces, and they have also been found to elicit increased amplitude of event-related potential (ERP) components as early as the P1 component, which peaks a mere 100-ms after stimulus onset. This entry describes the enhanced perception of emotional stimuli, effects of emotion on early perceptual processing, subliminal effects of emotional stimuli, questions about the function of preferential attention to emotional stimuli, and the influence of emotional states.

Enhanced Perception of Emotional Stimuli

Although early research emerging from the New Look movement gave consideration to both enhanced and impaired perception of emotional stimuli, recent work has tended to focus more on the former than the latter. Such work has found that people often are able to perceive emotional stimuli under conditions that would typically render non-emotional stimuli imperceptible. For example, when people search for two targets within a rapidly streaming sequence of stimuli, they often detect the first target at the expense of the second target, an effect known as the *attentional blink;* however, among neurally intact individuals, this effect is diminished when the second target is an emotional word. In other words, the emotional words seem to break through the attentional blink (patients with bilateral amygdala damage, however, show no such benefit for emotional words). Similarly, emotional stimuli tend to predominate over non-emotional stimuli and are perceived more readily during experimental manipulations in which separate inputs to the left and right eyes compete for awareness (e.g., binocular rivalry). Enhanced perception of emotional stimuli is not limited to the visual domain; for example, several experiments have employed "dichotic listening" tasks, in which people attend to auditory information presented to one ear while ignoring auditory information presented to the other ear. Although people often report being unaware of most information presented in the unattended channel, they do tend to notice particularly salient, emotionally meaningful words such as the sound of their own names.

Effects of Emotion on Early Perceptual Processing

In addition to neural evidence, behavioral evidence also suggests that effects of emotion on perception may occur early in the stream of processing. For example, one study found that presentation of an emotional face enhanced *contrast sensitivity* for subsequent targets. This term refers to the threshold of contrast at which people are able to disambiguate a stimulus from the background, and it typically is assumed to be an index of early vision. Although previous research has shown that contrast sensitivity tends to be enhanced via shifts of attention, evidence has also been found suggesting the effects of an emotional stimulus presentation and the effects of spatial attention on contrast sensitivity are additive, raising the possibility that emotion and attention might affect perception somewhat independently of each other. Although speculative, such possibilities converge with findings emerging from the neuroimaging literature: Although the impact of emotion on activity in visual areas can look similar to the impact of attention, some have suggested that the effects might possibly originate from and be instantiated via different neural pathways.

"Subliminal" Effects of Emotional Stimuli

Emotional stimuli tend to attract attention and engage some level of perceptual processing even when participants cannot report them because of laboratory-based manipulations. For example, in one study, high-contrast masks were presented to one eye and intact and scrambled versions of an erotic picture were presented to the other eye—a manipulation that, because of interocular suppression, leads the masks to dominate and obscure awareness of the pictures. Nevertheless, even though people could not report the presence of an erotic picture, they preferentially attended to its location, as revealed through higher target discrimination accuracy when a subsequent target appeared at the same location as the erotic picture rather than at the location of its scrambled counterpart. In the auditory domain, even when emotionally relevant words escape awareness in dichotic listening tasks, evidence suggests that they are perceptually processed at some level; for example, words that had previously been paired with electric shock—and that thus have some emotional significance—have been found to elicit indices of electrophysiological arousal, even when presented in the unattended stream and when participants report not having been aware of them.

Enhanced Perception as a Function of Preferential Attention to Emotional Stimuli?

Despite behavioral and neural evidence consistent with early effects of emotion on perception, the view that earliest stages of perceptual processing are open to the influence of emotion has been strongly challenged. Indeed, although the emotionality of a stimulus is reflected by activity in early sensory cortices, such findings provide relatively little insight into questions regarding where in the flow of information-processing emotion first exerts its influence. This is because it is difficult to determine whether such effects reflect modulation by emotion at initial processing stages or, instead, reflect "feedback" modulation instantiated via reciprocal projections from areas involved in higher-order processing. Indeed, notions that emotion (or any mechanism outside of a narrowly defined perceptual module) can affect early stages of perception are controversial. Although few would deny that the end product of perception can be colored by emotion, visual perception researchers such as Zenon Pylyshyn have argued forcefully that early stages of perceptual analysis—thought to extract and compute three-dimensional structure in the environment, among other visual tasks—must proceed beyond the influence of higher-order knowledge, expectations, or factors such as emotion and motivation; in other words, they must be cognitively impenetrable. According to this view, effects of emotion on visual processing could stem from effects on attentional allocation (which determines what receives perceptual processing in the first place) and on perceptual judgments and interpretations (which also contribute to our conscious experience), but most aspects of early visual computations lie beyond emotion's bounds. Even early appearing neural signals such as emotion-induced enhancements of the P1 ERP component resemble those that emerge simply as a function of attending versus ignoring a visual stimulus; thus, it could be that emotion influences perception solely through its impact on attention.

Many demonstrations of emotional influences on perception *can* be attributed to effects on attentional allocation. Emotional stimuli often seem to capture attention reflexively, and many of the manipulations that render non-emotional stimuli—but not emotional stimuli—imperceptible involve direct manipulations of attention. For example, the attentional blink described earlier is thought to render targets unreportable by presenting targets closer together in time than can be accommodated by the time it takes to make multiple attentional selections. Also consistent with the notion that emotional stimuli grab attentional resources that support perceptual awareness, the presence of emotional stimuli can actually impair perception of neighboring information: when people monitor a rapid stream of items for a single target, they often fail to detect the target when it is preceded by a particularly emotional distractor.

Influence of Emotional States

Although some research endeavors have focused on the role of emotional stimuli, others have focused on how emotional states influence perceptual processing in general. For example, experiments stemming from what is known as the *broaden-and-build* hypothesis—which suggests that positive emotion widens the scope of attention—have found that inductions of positive mood tend to bias participants to attend to "global" aspects of visual information rather than to "local" aspects: participants who underwent a positive mood induction showed an increased bias to make matching judgments based on global, rather than local, aspects of stimuli in which the local and global aspects contained conflicting information (e.g., where an assemblage of small squares (local) are arranged to form a larger triangle (global)). Such evidence that emotional states influence perception suggests an important bridge between perception research and clinical psychology, the latter of which focuses heavily on individual differences in chronic emotional states.

Steven B. Most

See also Attention: Covert; Attention: Effect on Perception; Attention and Consciousness; Attention and Emotion; Face Perception; Individual Differences in Perception; Pain: Cognitive and Contextual Influences

Further Readings

Alpers, G. W., & Gerdes, A. B. M. (2007). Here is looking at you: Emotional faces predominate in binocular rivalry. *Emotion, 7,* 495–506.

Gasper, K., & Clore, G. L. (2002). Attending to the big picture: Mood and global versus local processing of visual information. *Psychological Science, 13,* 34–40.

Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences USA, 103,* 17048–17052.

Most, S. B., Chun, M. M., Widders, D. M., & Zald, D. H. (2005). Attentional rubbernecking: Cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin & Review, 12,* 654–661.

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science, 17,* 292–299.

Vuilleumeier, P., & Driver, J. (2007). Modulation of visual processing by attention and emotion: Windows on causal interactions between human brain regions. *Philosophical Transactions of the Royal Society B, 362,* 837–855.

Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion, 7,* 882–886.

EVENT PERCEPTION

Event perception is the ability to recognize specific events, to recognize objects in events, and to perceive properties of those objects such as their shape and their weight as well as the specific ways that those objects move. The following is an illustration of event perception: A woman sets out for a walk in the park on a breezy fall day. She is able to see that she is walking and see the world go by as she walks. She can see the trees blowing in the breeze and falling leaves swirling about in the air. Squirrels can be seen leaping among the branches. She sees other people walking or jogging through the park. She sees a child swinging on a swing and another playing hopscotch while a third runs by while bouncing a ball. She sees an elderly man who is lifting a bag that she can see is heavy, so she offers the old man some assistance. All of these acts of perception entail event perception. This entry discusses the beginnings and significance of event perception, auditory and haptic event perception, the problem of information, event segmentation, and mirror neurons and motor theory.

Beginnings and Significance of Event Perception

The study of event perception was begun in 1950 by Gunnar Johannson, who, in his dissertation, actually started two closely related areas of investigation: event perception and structure-from-motion (or SFM). SFM is about seeing objects in events. One can see the rigid three-dimensional (3-D) shape of an object (is it a ball or an egg, a round can or a flat flask?) from seeing the motion of only a few random points on the object's surface. Johansson introduced a technique for isolating patterns of motion as visual information about events and objects in events. He distributed bright dots or patches on the surfaces of moving objects such as a person and filmed them in the dark so that only the dots showed up in the film appearing against a dark black background. When such movies are freeze-framed, so that only any single image from the movie can be seen, it just looks like a random distribution of dots. Without the motion, no three-dimensional structure can be seen. However, as soon as the movie is set in motion, it becomes obvious that it is a person walking.

At the first International Conference on Event Perception in 1981, one could witness about 200 of the leading perception scientists at the time gasp as they listened to Johansson and observed one of his now classic *point-light person* demonstrations. An unmoving arrangement of random dots suddenly moved and instantly could be seen to be a person sitting cross-legged and getting up out of a chair. The scientists likely gasped for two reasons. First, the informative power of patterns of motion was obvious. In such displays, you can see that it's a person and what he or she is doing. You can identify whether the person is male or female, how old the person is, whether the person is fatigued, whether the person is carrying a weight and if so, how much. You might be able to recognize a person's emotional state or who the person is. Established perceptual theory described information in static images formed when light was reflected from the entire surface of an object. But here there were just a few random dots in motion and not much, if anything at all, was gained perceptually by turning on the lights! It was all in the motion.

The second reason they gasped is because of what this implied about the role of time in perception and experience. Perceptual scientists were used to thinking about the information presented to the nerves in the retina on the back of the eye by a single momentary static image that was then processed by the brain to derive a perception of objects in the world. Event perception meant that the information as well as the thing seen were both distributed over significant lengths of time. The perception of a bouncing ball in a point-light display requires the time for the ball to fall, hit the ground, and rebound to provide enough spatial-temporal structure to enable one to recognize the event. This takes seconds, not milliseconds. Scientists still do not well understand how the brain might work with this kind of spatial-temporal information.

Auditory Event Perception

Event perception involves sensory modalities in addition to vision like audition (that is, hearing) or haptics (that is, touch and kinesthesis). We can recognize events from the sounds they make. One can recognize a bouncing glass bottle that has been dropped and distinguish it from a plastic one or one made of glass that shatters. One can tell from the sound of the water being poured into a glass when the glass has become full! Of course, one of the most puzzling feats of acoustic event perception is the ability to hear what someone is saying, that is, speech perception. The production of speech is an event involving movements of the vocal bands, the tongue, the lips, and the teeth of the speaker. Speech perception is like the visual perception of point-light people in that one can recognize, for instance, familiar people or the gender of a person or their emotional state from their speech as well as from what a person is saying.

Historically, the information in sound for speech perception has been analyzed by trying to make it as much like static snapshots as possible. The stream of sound has been analyzed into extremely brief elementary sounds such as letters and phonemes. The problem with this is that the phonemes change radically in the context of a continuous stream of speech where a given phoneme is preceded and followed by other phonemes. They interact in the production (acoustically, the "t" at the beginning of "titter" is entirely different from the "t's" in the middle of the word), but this does not seem to hinder one's ability to perceive what is being said. This problem is called *co-articulation.* The implication is that a more time-extended aspect of the sound patterns of speech may actually carry the information.

Haptic Event Perception

Objects, both animate (a cat versus a hamster) or inanimate (a potato masher versus a whisk), can be recognized by touch or haptic perception (with eyes and ears closed). This is a form of event perception. The reason is that people cannot readily recognize things just pressed into the passive palm and fingers of their hand. If they are allowed to manipulate an object in the hand actively, then recognition is rapid and accurate. The spatial-temporal structure generated as the hands interact with the surface textures and surface compliances (rigid, elastic, plastic) and surface shapes of objects yields information about those objects. Speech can be recognized haptically. Some people who are both blind and deaf can perceive what a person is saying by placing their open hands over the bottom of the speaker's face covering the lips and the jaw with a thumb on the throat touching the voice box or larynx. This is called the Tadoma method.

The Problem of Information

What is the information that allows one to recognize events? What is it about the spatial-temporal pattern of moving dots in a point-light people display or a structure-from-motion display that allows one to recognize the event or the object? Attempts to answer these questions can be categorized into two types: one analyzes kinematics and the other analyzes dynamics. *Kinematics* is the motions in events. Kinematic properties include time or frequency, position and its change over time, velocity and acceleration and their changes over time, and so on. Kinematics is lengths and times. *Dynamics* is about mass and the force or energy required to move it. *Dynamics* is concerned with the causes of motion. It entails the Newtonian laws of motion that generate the specific forms of events: pendulums, projectile motion, bouncing balls, and yes, walking people.

Johansson's approach was kinematic and he called it perceptual vector analysis. He analyzed the relative and common motions of the dots in the display. His point was that we really see motions in these ways. So, for instance, if you have a dot moving horizontally along the bottom of (an invisible) square and another dot moving at the same time vertically along the right edge of that square so that the two dots met at the bottom right corner and then reach the bottom left corner and the top right corner, respectively, at the same time (that is, the two dots are moving in phase), then what a person sees is not the motions as they have been described so far. Instead, what one sees is the two dots moving directly toward one another (as they approach the bottom right corner) and then directly away from one another (as they move to the bottom left and top right, respectively), and the line along which they are perceived to move itself moves diagonally across the square (that is, from the bottom right to the top left). The perceived motion of the dots toward and away from one another along the line is *relative motion* and the diagonal motion of the entire "line" is *common motion.* The discovery of this principle was important because it meant that one does not see the motion of the dot on the foot of a point-light walker as a hopping motion along the ground, but instead, as a pendular motion relative to the dot on the hip while the whole person moves over the ground. Johansson was a Gestaltist, and this analysis was characteristic of Gestalt. This kind of analysis was elaborated into an entire field of research on structure-from-motion (SFM)

James Todd performed a study using the kinematic approach in which he concluded that the approach could not be used to solve the problem. He set out to investigate how observers distinguish and recognize two different human gaits, namely, walking and running. Todd recorded the actual human motions in each case and then made stick figure animations of legs walking and running. He analyzed each into motion functions for each joint, and then he mixed the joint motions to try to determine which was the most important for making the whole gait look like either walking or running. The result was that none of them did. As he increasingly mixed the motions at different joints, the resulting gaits simply became more ambiguous. Basically, his approach was to perturb the motions in an event in an effort to discover what it was about the motions that made them recognizable as what they were. At the end of his study, Todd concluded that the space of possible motions was much too vast to be able to investigate in this way.

Subsequently, researchers in biomechanics revealed that walking and running are distinguished dynamically. Walking consists of upright pendulums and an inverted pendulum. The leg with the foot on the ground (the stance leg) is an inverted pendulum (like an old fashion metronome). The other leg (the swing leg) is a coupled set of upright pendulums. Two legged robots have been built that walk in this way with no motors down a slight slope, and they move and look just like a human walking. Running consists of a bouncing ball or pogo stick dynamic (elastic rebound and projectile motion). Running robots have been built using these dynamics and again, they look alive.

In a dynamic approach, events are characterized by the underlying dynamics (physical laws) that generate the specific forms of motion characteristic of a given recognizable type of event. In this approach, events are recognized like objects, namely, by their shapes. Object shapes are spatial. Event shapes are spatial-temporal. Each point moves along a 3-D path of specific shape and as it does so, the velocity of motion varies in a specific way. The result is what has been referred to as a trajectory form. A specific dynamic generates a specific trajectory form. Geoffrey Bingham has shown that human observers can discriminate among a variety of different trajectory forms and are able to use them to recognize different events. Furthermore, they were able to do this reliably when they saw the events from different perspectives.

Gravitationally governed events, such as a walking person or a bouncing ball, that have been inverted in a point-light display are difficult to see and to recognize. Thomas Shipley has shown that this is not simply because we are not used to seeing a person upside down, but instead because the motions have been perturbed with respect to the gravitation frame of reference. Shipley made point-light displays of a person walking on his hands (that is, inverted but consistent with the gravitational constraints on the motion). Observers had no difficulty seeing and recognizing this event.

So, both the study of the phase relations among moving points (that is, the timing of relative motions) and dynamically generated trajectory forms provides some insight into the visual information that allows people to recognize events. However, it remains difficult to discover what it is, for instance, about the walking motions of men and women that enables people to recognize the gender of point-light walkers. Nikolaus Troje used a kinematic approach to provide some understanding of this. He used an advanced correlational method (principal component analysis or PCA) to analyze the motions in point-light displays that were identified by observers as being male or female walking. In this way, he isolated properties that he could vary to turn a point-light display into either male or female walking. He was able to manipulate the motion to make it look one way or the other. This analysis, however, did not reveal what generates the specific recognizable forms of motion.

Dynamical analysis has revealed some of the information that allows us to see how big things in events are. Motions governed by gravity yield timing that covaries with spatial scale. For instance, the period of a pendulum's swing is a function of its length. The legs of animals, including humans, are essentially pendulums. So, the timing of their motions provides information about their scale, and human observers have been shown to be sensitive to this information.

Event Segmentation

Events are continuous in time and some events, especially those involving human actions, are so complex that they are composed of both a hierarchy and a series of simpler events. This was true of the walk in the park described earlier or, for instance, the act of cooking a meal. A meal might involve multiple dishes to be prepared, and the preparation of each dish involves a series of tasks, each of which involves subtasks such as reaching to grasp, stirring, peeling, or dicing. The whole business runs continuously in time potentially against a background of other distinct, but similarly continuous complex events (such as kids doing their homework at the kitchen table). Thus, perception of the event requires that it be segmented both from its background and into its various component parts. Barbara Tversky and Jeff Sacks have addressed this problem by showing observers movies of complex events and asking them to indicate where the simpler component events start and stop. Tversky and Sacks find that observers are consistent in their segmentation of the events. Their theory is that observers have mental representations of events that are used to predict what should happen in the immediate future and that these predictions are continually compared with what actually happens. When one event ends and another begins, the predictions momentarily become less accurate because, until the appropriate event representation can be identified, the course of motion cannot be anticipated. The idea is that observers use error rates in their predictions to determine boundaries between events. Many aspects of the theory have not yet been tested, but it's all fairly new.

Mirror Neurons and Motor Theory

Finally and more recently, there has been much excitement generated among researchers by the discovery of mirror neurons. These are neurons found in monkey brains that light up when the monkey sees the goal-directed movement of either its own or another monkey's hand reaching for something. This has led to a body of research pursuing the idea that the perception of biological motion (that is, point-light walkers rather than bouncing balls) is special. This theory is reminiscent of a motor theory of speech perception that was developed at Haskins Laboratories in the 1970s. The idea is that these events are seen in terms of the perceiver's own motor system. A limitation of this theory is that it cannot explain how events other than biological motions are perceived.

Geoffrey P. Bingham

See also Audition; Embodied Perception; Gestalt Approach; Haptics; Kinesthesia; Mirror Neurons; Motion Perception; Motion Perception: Physiological; Speech Perception

Further Readings

Epstein, W., & Rogers, S. (1995). *Perception of space and motion.* San Diego, CA: Academic Press.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events.* Cambridge: MIT Press.

Jansson, G., Bergström, S. S., & Epstein, W. (1994). *Perceiving events and objects.* Hillsdale, NJ: Lawrence Erlbaum.

Port, R. F., & van Gelder, T. (1995). *Mind as motion: Explorations in the dynamics of cognition.* Cambridge: MIT Press.

Shipley, T. F., & Zacks, J. M. (2008). *Understanding events: From perception to action.* Oxford, UK: Oxford University Press.

Warren, W. H., & Shaw, R. E. (1985). *Persistence and change: Proceedings of the First International Conference on Event Perception.* Hillsdale, NJ: Lawrence Erlbaum.

EVOKED POTENTIAL: AUDITION

Auditory evoked potentials (AEPs) represent synchronous neural activity in the brain in response to a sound changing—for example, in frequency or level—and that can be recorded from electrodes attached to the scalp. AEPs can be used to detect hearing loss in newborns and infants; they are exquisitely sensitive at the millisecond (ms) level to changes in conduction time in axons and temporal processing in neurons. This high time resolution allows the tracking of maturational changes in the brain, the documentation of short-term auditory learning, and the diagnosis of disorders related to temporal processing in general.

AEPs comprise activity from neural generators in the auditory nerve, the fiber tracts in the brain stem, and the pyramidal cells in the auditory cortex. Table 1 provides an overview of type, origin, and use of the response components. This entry discusses compound action potentials and compound synaptic potentials, the auditory brain stem response, obligatory cortical evoked potentials, and auditory memory.

Compound Action Potentials and Compound Synaptic Potentials

The neural activity underlying AEPs, typically evoked by clicks or short tone pips, consists either of compound action potentials, that is, composed of firings occurring at the same time, in nerve tracts that make up the auditory brain stem response (ABR), or of simultaneously occurring postsynaptic potentials in cortical cell dendrites that make up the cortically generated evoked potentials. Action potentials are of much shorter duration (about 1 ms) than synaptic potentials (about 15 ms) and because the extracellularly recorded activity is biphasic, action potentials can cancel each other for small time differences whereas postsynaptic potentials typically are much less sensitive to this. In both cases, the amplitude of the AEP components is proportional to the number of synchronously activated nerve fibers (in case of the ABR) or number of synchronously activated synapses on cortical pyramidal cell dendrites. AEPs can only be detected if the generating elements (nerve fibers, dendrites) are spatially aligned and currents therein flow in the same direction.

The Auditory Brain Stem Response

The ABR is a short latency subset of the AEPs. It is a multi-peaked response that alternates between vertex positive and negative voltages starting with a positive peak termed *wave I,* followed with 1-ms interval by further positive peaks II and III, and then by an often double-peaked wave IV/V. Waves I and II originate from the extracranial and intracranial part of the auditory nerve, respectively, wave III from the first auditory nucleus in the brain stem, and wave IV/V from the tract of the lateral lemniscus on its way to the auditory midbrain. The ABR threshold matches that of behavioral measures. The amplitude at the scalp is small (about 0.1 mV). The latency difference between waves I and V represents the brain stem conduction time, including synaptic delays and nerve conduction.

Table 1 Auditory Evoked Potentials: Type, Origin, and Use

Name	*Response latency*	*Origin*	*Generator*	*Use*
Auditory Brainstem Response (ABR)	1–10 ms	Auditory nerve and nerve tracts in the brainstem	Synchronized action potentials	• Newborn hearing screening • Disorders affecting nerve conduction
Middle Latency Response (MLR)	18–50 ms	Primary auditory cortex on Heschl's Gyrus	Synchronized post-synaptic potentials in pyramidal cell dendrites	• Cortical maturation • Frequency-place map changes • Learning
Long Latency Obligatory AEPs	5250 ms	Association cortex on the Planum Temporale		
Mismatch Negativity (MMN)	150–300 ms	Superficial layers in primary auditory cortex		• Pre-attentional sound, melody and phoneme discrimination

This latency difference is a sensitive indicator of auditory brain stem maturation in the first year of life and of the presence of tumors pressing on the auditory nerve or brain stem, diseases such as multiple sclerosis that slow nerve conduction, and auditory neuropathy. Immaturity and the listed disorders all increase the brain stem conduction time or make waves beyond wave I undetectable or, as in auditory neuropathy, do not show any wave in the ABR. The ABR is unaffected by sleep and light sedation or anesthesia and is therefore of prime interest in newborn hearing screening.

Obligatory Cortical Evoked Potentials

An oscillatory complex with period of about 25 ms is found in the latency range of 15 to 50 ms and is called the middle latency response. When the stimulus is presented with intervals around 25 ms (about 40 hertz [Hz]), a steady-state response mimicking a 40-Hz sine wave results. The middle latency response is composed of activity entering and leaving the primary auditory cortical areas situated on Heschl's gyrus.

A slower oscillation with a period of about 100 ms reflects activity in secondary cortical areas situated on the planum temporale. This oscillation is characterized by a positive (P)–negative (N) P1–N1–P2–N2 sequence with approximate latencies of 50, 100, 150, and 220 ms. Although the amplitude of these components is affected by sleep, sedation, and attention, they are always present in normal hearing adults. The dependence on attention makes these responses not suited for threshold estimation. The P2 component is the first to mature (about 2 years of age) whereas the N1 component matures to well into the late teens. The N1–P2 complex increases in amplitude as a result of learning an auditory task on a time scale of a few hours.

Auditory Memory

When a sequence of tone pips with frequency of 1 kilohertz (kHz) and presented at a rate of about 1 per second is randomly interrupted by a tone pip with different frequency (say, 1,100 Hz, called the oddball), the obligatory response to the oddball is different from that to the frequent stimulus just preceding the oddball. The difference between the response to the frequent and oddball stimuli represents a dominant negative wave peaking about 150 ms latency. This is called the mismatch negativity (MMN). It does not require attention and is used to objectivate the ability of the person to detect acoustical differences such as frequency, duration, and level, but also phonemic and melodic differences. When the interval between the tone pips becomes larger, the size of the MMN decreases and altogether disappears around a 10-second interval between the stimuli. This has been equated with the length of the auditory memory span.

The cortically generated AEPs are complementary to imaging methods such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) that reflect local energy requirements resulting from neural activity. The AEPs have an exquisite temporal resolution and reflect the degree of neural synchrony in the activated neural population. PET and fMRI only depend on the number of activated neurons in the population not on the synchrony of their activity. In contrast, the spatial resolution of active neural populations is better for the imaging methods, especially fMRI.

Jos J. Eggermont

See also Audiology; Auditory Thresholds; Evoked Potential: Vision; Magnetoencephalography

Further Readings

Burkard, R. R., Don, M., & Eggermont, J. J. (Eds.). (2006). *Auditory evoked potentials: Basic principles and clinical application.* Baltimore: Lippincott Williams & Wilkins.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118,* 2544–2590.

Picton, T. W., & Taylor, M. J. (2007). Electrophysiological evaluation of human brain development. *Developmental Neuropsychology, 31,* 249–278.

EVOKED POTENTIAL: VISION

A *visual evoked potential* (VEP) refers to a systematic change in the electrical activity of someone's brain in response to a change in what the person is viewing. It is revealed by averaging over many trials the raw electrical activity, the electroencephalogram (EEG), recorded from 2 to 256 electrodes at standard positions on the person's head from each individual trial. For example, if a person is looking at a set of lines with one eye and at an oppositely oriented set of lines with the other eye, and then one set of lines changes to be identical to the other, the EEG from electrodes on the back of the head changes in a typical way over about 400 milliseconds (ms): a VEP (color insert, Figure 23a). The VEP is an event-related potential (ERP), because it is linked to a specific stimulus event—in this case, a visual stimulus. VEPs are used for clinical purposes, such as for locating problems to particular parts of the visual system, and for research. This entry describes VEPs, their advantages and disadvantages, and their clinical and research uses.

From the EEG to the VEP

In 1929, the German psychiatrist and neurologist Hans Berger placed an electrode on a person's scalp, recording the first EEG. He discovered that when the person closed his or her eyes, the EEG altered reliably even on a single trial, anticipating the VEP. In 1939, Pauline Davis reported that a unique, reproducible deflection of EEG activity could be evoked by an auditory stimulus. This was the first evoked potential, visible on a single trial. In 1960, Bill Cobb and George Dawson recorded the first, unambiguous VEP, averaged over many trials.

The fluctuations in electrical activity of a VEP depend on the location of the electrodes and on the sort of visual change. The VEP in the color insert, Figure 32(a), shows its first major component as a positive wave around 100 ms after the change, the P100 or P1. After that is a negative wave around 175 ms (N175, N2).

The amplitude of any VEP, peak-to-trough, is around 10 microvolts (μV), much below the spontaneous brain electrical activity of about 100 μV, so to see a VEP one needs to repeat the visual change about 100 times and average the EEG. This reveals the VEP, time-locked to the change, whereas the other activity, not time-locked, is averaged out (color insert, Figure 23a).

No one is sure what generates VEPs, but for many researchers, the most convincing theory is that VEPs arise from postsynaptic potentials (PSPs) over the dendritic trees and cell bodies of neurons. What is certain is that VEPs are not from action potentials, which are too brief. In any individual neuron, PSPs create a *dipole,* a gradient of electrical activity from positive (e.g., cell body for excitatory PSPs) to negative (dendrites). When a huge number of neighboring neurons are parallel to each other and have the same dipole for more than a few milliseconds, this can be recorded from an electrode on the scalp. The P100, therefore, represents simultaneous dipoles from one set of neurons; the N175 represents simultaneous dipoles either from a different set of neurons at a different angle or direction to the electrode or from inhibitory PSPs in the same neurons.

A crude measure of where in the brain a VEP's dipoles (its sources) are located is which electrodes show it most clearly (see color insert, Figure 23b). More precise measures try to reconstruct the dipoles that gave rise to the voltages in electrodes scattered over the scalp. In theory, any distribution of electrode voltages can be produced by an infinite number of one or more dipoles. In practice, constraints on the number of dipoles (e.g., fewer rather than more) and on their location (e.g., in grey matter rather than in white), yield solutions having a spatial resolution below 10 cubic millimeters (mm^3). One such approach is illustrated in Figure 23(c) in the color insert.

Advantages and Disadvantages

VEPs have at least three advantages over other methods of measuring electrical activity in the brain. First, VEPs are inexpensive to obtain, much cheaper than functional magnetic resonance imaging (fMRI) or magnetoencephalography (MEG). Second, VEPs are noninvasive. This differs from techniques such as positron emission tomography (PET) (requiring observers to ingest radioactive material) and local field potentials (requiring observers to have electrodes inserted into their brains). Third, VEPs have temporal resolution in the millisecond range. This is about 1,000 times better than PET and fMRI. VEPs' only disadvantage is their poorer spatial resolution (about three times coarser than PET and fMRI).

Clinical Uses

A corneal electrode and brief (< 1 ms) flashes of a homogeneous field, *flash stimulation,* yields the electroretinogram (ERG). This indicates rod or cone function in the *a*-wave at about 20 ms and bipolar function in the *b*-wave at about 50 ms. A pattern such as a checkerboard, with the black squares turning white and the white squares turning black at 2 Hz, gives *pattern (reversal) stimulation.* This evokes the pattern electroretinogram (PERG), from the retinal ganglion cells.

Scalp electrodes over the occipital lobe with flash and pattern stimulation yield the *flash VEP* and the *pattern (reversal) VEP* respectively. These assess the entire visual pathway up to V1, with pattern VEPs being more sensitive. The finest pattern VEP is correlated with visual acuity. Characteristics of pattern VEPs are well established in ophthalmology and neurology where diagnostic uses include finding lesions that block transmission along the visual pathway, revealing pathway misrouting (e.g., in albinism), and identifying malingering (people falsely claiming to be blind).

A typical VEP to pattern reversal has a N75, a P100, and a more variable N135 (N2). To pattern onset stimulation, many display a P90 (C1), an N120 (C2), and a P200 (C3). Amplitude of components varies by more than a factor of two among people.

Research Uses

Just about every aspect of visual perception has been researched using VEPs. All that is required is some visual event that can be specified with millisecond accuracy. Manfred Fahle and Michael Bach recently reviewed the wide range of research topics, which include the effects of visual attention, visual memory, and type of visual stimulus. With its high temporal resolution and its reasonable spatial resolution, VEPs offer an invaluable perspective on brain activity associated with visual perception.

Robert P. O'Shea, Urte Roeber,
and Michael Bach

See also Brain Imaging; Consciousness; Evoked Potential: Audition; Magnetoencephalography; Oscillatory Synchrony

Further Readings

Cobb, W. A., & Dawson, G. D. (1960). The latency and form in man of the occipital potentials evoked by bright flashes. *Journal of Physiology, 152,* 108–121.

Fahle, M., & Bach, M. (2006). Basics of the VEP. In J. Heckenlively & G. Arden (Eds.), *Principles and practice of clinical electrophysiology of vision* (pp. 207–234). Cambridge: MIT Press.

Luck, S. J. (2005). *An introduction to the event-related potential technique.* Cambridge: MIT Press.

EVOLUTIONARY APPROACH

Perception is a universal phenomenon. It functions primarily as a means of allowing an organism to process changes in its external environment. Thus, perception has substantial survival value and can be observed in all living species. Its prevalence leads to questions about its origins and about how perceptual mechanisms have evolved. These questions require a firm understanding of evolution, particularly the main processes of innovation and constraint, and their interrelationship.

Evolution is considered as change in the inherited traits (particular characteristics) of a population of organisms from one generation to the next. These changes are caused by a combination of three main processes: variation (differences between genotypes underlying the traits, caused by, for example, mutation), reproduction (heritable traits are propagated between generations), and selection (individuals with advantageous traits are more likely to reproduce). In the 150 years since Charles Darwin published *The Origin of the Species,* scientists have gained an understanding of a great number of biological phenomena that could not have been achieved without a firm knowledge of the mechanisms of evolution. Perception is one of these.

Perception using an evolutionary approach allows us to search for the causal explanations of the adaptations that are seen in modern-day organisms. All traits have both mechanical and evolutionary causes, and a complete explanation of perception requires an understanding of both of these, as described in this entry.

Questions About Evolution

A complete understanding of an organismic trait such as perception requires answers to four main questions. For example, the question "Why does a pigeon see five colors?" can mean four different things, some times referred to as *Tinbergen's four whys* after the pioneering ethologist Niko Tinbergen. The different meanings of "Why?" originate from different levels of explanation. The first "why" examines the function of a trait, asking, for example "How does the eye work?" This is a question of (proximate) causation, studied mainly by (neuro-) physiologists. The second question is about the survival value of a trait; for instance, why is visual perception useful? The third question asks how an ability arises during an individual's development (ontogeny) given some early experiences and the genetic makeup. And finally, it is important to know how a particular behavior evolved, how it came into existence. This requires inference about the behavior of ancestral species, examining how and why this behavior changed over generations.

In the animal kingdom, there is not one brain, one sense, one problem, and one solution. At every level of organization, variation prevails. Of great value to understanding how novel features emerge is the study of the relationships between individual development and phenotypic change during evolution (the new science of Evo-Devo). The inescapable difficulty is tracking the process of innovation and variation (resulting in the evolution of new traits) on the one hand, and stabilization and conservatism (resulting in traits remaining stable across time) on the other. Why have many solutions to the same problem evolved in some cases, but in other cases only one? Why did some structures not disappear despite their obvious malfunction (e.g., the appendix)? Why are the receptor cells in the retina facing backward in vertebrates but not in invertebrates?

Answers to these questions demand a rigorous evolutionary and comparative approach. In a nutshell, the evolutionary paradigm holds that the result of natural selection is the generation of adaptive organic structures that are sustained as the result of two factors: correspondence to external circumstances and coherence between the components of the system. The latter follows the principle of organization and is particularly important for such complex organs as the eye, in which many elements must work together. Therefore, explaining how animals solve the sorts of problems they face in their natural habitats cannot be achieved by applying a purely functional (and sometimes anthropomorphic) approach.

Some people think that perceptual faculties (e.g., perceiving color) are the result of selective forces exerted by specific environmental circumstances (e.g., discriminating between flowers, or between ripe and immature fruits). However, evolution is seldom as straightforward as this. The adequate metaphor is to view evolution as a "tinkerer," building innovations by modifying existing structures, rather than by designing them from scratch. A famous example is the transformation of jaws into the three ear bones (the hammer, anvil, and stirrup). Understanding the main classes of evolutionary processes—innovation and constraint—and their interrelationship, requires a firm understanding of evolution and the adoption of a systemic viewpoint.

An important question asks how major changes evolve and how new features originate. First, we can ask what the genetic and developmental bases of such changes are. Second, we can ask what role natural selection plays in their evolution. For instance, we may ask whether each step, from the slightest initial alteration of a feature to the full complexity of form displayed by later descendants, could have been guided by selection. A major problem here is to understand what functional advantage there can be in an organ that is not completely developed. The evolution of the eye is a good example of this. Darwin himself was skeptical about whether natural selection could explain a stepwise progression to a fully functional vertebrate eye. In particular, one must question how such a complex organ could have evolved if its proper function depends on the mutually adjusted forms of each of its components (e.g., lens, cornea, retina, blood vessels).

The Evolution of the Eye

Organisms have exploited light for perceptual information through the evolution of photoreceptors and, ultimately, eyes. However, only a handful of eye types exist. One reason for this is that the physics of light constrains photodetection; the other is the result of the constraints of evolution.

Visual perception was well established by the time there is any fossil record of life on Earth. Most basic eye types arose in a brief time during the Cambrian period, around 530 million years ago. As soon as organisms became more mobile and predation became a common way of life, more efficient eyes developed, and they developed rapidly. Computer simulations have revealed that the evolution from a patch of shadowed photosensitive tissue to an eye resembling that of a fish could take as little as half a million years.

Only about a third of the animal phyla have no specialized organ for detecting light, a third have light-sensitive organs, and the rest possess eyes. Image-forming eyes appeared in 6 of the 33 extant metazoan phyla (Cnidaria, Mollusca, Annelida, Onychophora, Arthropoda, and Chordata); these six contribute to about 96% of the known species alive today. This simple calculation shows that the evolution of eyes (and thus visual perception) contributes significantly to survival.

It is thought that the eye has evolved independently about 40 different times. However, there is much structural similarity between different eyes—for example, there are few types of photoreceptor. Photoreceptors might predate eyes and might have evolved from a single common root; however, it is also feasible that, because of structural constraints, they represent the only possible solutions. Furthermore, the photopigment rhodopsin appears in all animals, and the master control gene (Pax6), which causes eye development in the embryo, is amazingly similar in all vertebrates and invertebrates (that have been investigated so far). This apparently suggests that all eyes have a common ancestry. However, we do not know if the master control gene originally evolved to control the development of an eye with spatial vision, or a directional photoreceptor, or perhaps just the simplest form of light sensitivity. Furthermore, despite the discovery of profound homologies (similarity by common descent) in the regulatory genomes of a diverse set of organisms, other levels of homology in the evolution of the eye must be considered. It is likely that the photopigment molecule originated a long time ago in a common ancestor to all animals. Photoreceptor cells then evolved independently a handful of times and spatial vision (real eyes), numerous times, all within the physical constraints of the preexisting structure.

Ecology and Evolution

We can see the impact of the environment on the evolution of the eye by looking at animal species that are alive today. Different visual abilities can be observed even within a species. New world primates exhibit an unusual variation in color vision, all the males have dichromatic vision, but the females can either be dichromats or trichromats. This difference is thought to be maintained by selective forces that act differently on the two morphs. Trichromats are better at finding ripe fruit, whereas dichromats are better at detecting objects with red-green camouflage (such as insects). Natural selection plays a key role in maintaining this polymorphism.

Eyes have also become redundant as a result of changes in selective forces. Cave animals are ideal subjects for the study of evolutionary change, and are particularly important for perception. Examining how an animal that lives in complete darkness has adapted from its surface living ancestor can reveal much about the key factors that play a role in the evolution of perception. Mexican cavefish (tetra) have evolved in multiple isolated populations. Interestingly, the impact of the change in environment has revealed similar traits in the different populations. The darkness of the caves leads to limitations in food, space, and reproduction; these have led to a specific suite of perceptual adaptations in the cavefish. They have evolved increased chemical senses and specialized sensory organs that respond to mechanical input such as tension, pressure, or displacement (mechanoreception). However, the cavefish still have small, nonfunctioning eyes. The eyes have not entirely disappeared, so it is likely that they no longer hinder the animal and thus there is no evolutionary pressure to reduce them further.

Interestingly, the auditory and olfactory systems have remained largely unchanged in the cave dwelling populations. The consequence of living in caves has had different effects on the perceptual abilities that can be seen in the surface dwelling populations. Those senses that still work in a cave environment (olfaction) remain the same; however, those that are no longer useful (eyes) regress, and novel perceptual mechanisms (mechanoreception) appear.

Origin of Perception

A major task for evolutionary biologists is to track the origins of traits. When one thinks of perception, it tends to be in terms of the visual; humans (and other primates) are highly visual animals. However, this level of sophistication is not available to single-cell organisms. For a population of organisms to survive, the individuals within that population must be able to recognize a suitable living environment (and avoid poor environments), recognize food items, and recognize mates. For an organism to fulfill these tasks, it must have the ability to survey the world outside its own body and respond to the information that it receives. That is, the organism must be able to perceive its surroundings (in some way) and act upon them.

If the evolutionary pressure to perceive is so great, then we would expect that even single-celled organisms should be able to perceive at a very basic level. One does not normally think of bacteria or protoctista as having sensory perception or even possessing behavior for that matter. However, they possess a rich repertoire of responses (changes in the metabolism of the cell, changes in movement direction, production of molecules that are perceived by neighboring organisms), which suggests that the basic mechanisms to perceive our environment evolved when most organisms were but single cells. By using surface receptor molecules, single-cell organisms are able to distinguish between a wide range of different molecules. Study of these very simple organisms and their perceptual mechanisms can provide essential evidence for uncovering how perception evolved in its earliest forms. Further, it elucidates the basic "systemic" principle of closing the loop from perception to action and again to (new) perception.

Processing

There are two possible ways of processing sensory information. Senses can be processed in parallel, with information from different senses being transmitted along separate lines. They can also be processed centrally, in which case a neuron's receptive field (the area on the receptor surface that influences firing of the neuron) is modified by synapses in the central nervous system (CNS) and processed together in the sensory cortex, resulting in interaction across senses. The parallel organization is quite different and would result in a representative system for each sense organ (eyes, ears, nose, etc.), which allows the possibility that an animal would live in several different sense worlds simultaneously. Reptiles possess this processing system. For example, snakes use different senses separately when hunting. The snake uses visual perception to track and intercept the prey, once struck, it uses only its sense of smell to locate the prey (it is particularly interesting that boas and pythons do this even when the prey is held in their coils). Then the snake searches for the prey's head (they swallow food head first), and for this, only the sense of touch is used.

This example reveals how an organism can successfully use a number of separate systems to hunt and feed. However, an entirely centralized system (as can be seen in mammals) is able to combine the information received from the sense organs and their associated brain areas. It thus generates a multimodal representation of the world; this evolutionary breakthrough allows much greater possibilities for cognitive representation and behavioral flexibility

From Perception to Cognition

The primary function of the brain is to compute dynamic, predictive models of the environment. Across the animal kingdom, organisms are able to rapidly evaluate their current situation and respond appropriately to it. This suggests that the perceptual constructions of the external world provide meaning or functional significance to objects and situations. As humans, we perceive an object as having a particular shape or color, and we perceive it as a dog, or a tree (or whatever it is). Being able to identify objects as members of known categories allows the organism to respond to them in appropriate ways. Categorization can thus be seen as the end point of perception and the groundwork of cognition. Without the ability to perceptually classify items into known categories, it would be difficult to behave appropriately toward the enormous variety of new objects that are encountered daily. Classification allows a reduction of the vast amount of information that is perceived and takes advantage of the fact that objects within the same category share many properties and behaviors.

The evolutionary pressures to minimize processing requirements have led to several different solutions, ranging from selective attention to feature learning and abstraction. These abilities have been studied in a wide range of organisms, with such success that leading researchers in the field have suggested that evidence of categorization has been observed in every level of the animal kingdom in which it has been adequately sought. Pigeons, for example, are able to categorize pictures of both complex natural scenes and human-made stimuli. The underlying cognitive mechanisms include memorization of pictures, feature learning and prototype formation. All of these are based on the animal detecting the presence or absence of physical features. However, evidence indicates that animals are able to do more than this; they can classify stimuli based on abstract relational information.

"Sameness" is a good example of a relational concept. A very basic test of this would present an animal with a red circle, and then subsequently with a green circle and a blue square. If animals are able to categorize based on choice of the same shape, they should select the green circle because it is the same shape as the sample stimulus. However, it appears that access to knowledge about the relations between stimuli is easier for some species than for others. For example, primates, corvids, and dolphins can extract abstract rules more readily than the pigeon can; however, in many cases, pigeons are able to do this, but they have difficulty and are only successful if the experimenter makes it impossible for them to use any consistent visual cues or rote learning of specific stimulus arrays. This difference does not necessarily reflect a vast gap in cognitive and perceptual abilities, but suggests that the abstract rules are less salient to the pigeon than to (say) corvids. This is likely to be related to complexity of social structure. Further evidence that supports this idea can be seen in lemurs. Different species have different levels of ability to infer relations between objects; this appears to be related to differences in their social structure, with the social species being more readily able to perform at high levels on these tasks than are their solitary counterparts. To be able to judge complex relations between stimuli is the end point of perception and the start of cognition.

Cognition, which includes learning and memory, has evolved as a means of guiding behavior. However, the path from perception to action isn't necessarily long and indirect. We now know that the perceptual and the motor system are strongly interlinked from very early stages of processing. The mirror neurons in the premotor cortex F5 and the anterior intraparietal (AIP) area, which respond both when an organism carries out an action and when it observes another organism carrying out the same action, have both sensory and motor properties. They contribute to the sensory-motor transformations that are necessary to select and locate objects and implement the movements required to act on these objects. The AIP-F5 circuits categorize objects by possible actions that can be made upon them; for example, a cup can be picked up by the handle or by the main body. Upon perceiving an object (such as the cup), these neurons immediately code a given set of possible actions. Thus, they respond selectively to the objects by their affordances (action possibilities) and the potential motor acts that can be conducted on them. Most evidence for the presence of these neurons has been found in primates (macaque monkeys and humans). However, neurons with this mirror function have recently been discovered in songbirds. What is currently unclear is whether they are a very old and widespread phenomenon or the result of convergent evolution of perception-action processes.

Ludwig Huber and Anna Wilkinson

See also Animal Eyes; Evolutionary Approach: Perceptual Adaptations; Eyes: Evolution of; Visual System: Evolution of

Further Readings

Fernald, R. D. (2006). Casting a genetic light on the evolution of eyes. *Science, 313,* 1914–1918.

Futuyma, D. J. (2005). *Evolution.* Sunderland, MA: Sinauer Associates.

Jeffrey, W. R. (2001). Cavefish as a model system in evolutionary developmental biology. *Developmental Biology, 231,* 1–12.

Land, M. F., & Nilsson, D. E. (2002). *Animal eyes.* Oxford, UK: Oxford University Press.

Levine, M. W. (2000). *Fundamentals of sensation and perception.* Oxford, UK: Oxford University Press.

Müller, G. B. (2007). Evo–devo: Extending the evolutionary synthesis. *Nature Reviews Genetics, 8,* 943–949.

Nilsson, D.-E., & Pelger, S. (1994). A pessimistic estimate of the time required for an eye to evolve. *Proceedings of the Royal Society of London B, 256,* 53–58.

Surridge, A. K., Suarez, S. S., Buchanan-Smith, H. M., & Mundy, N. I. (2005). Non-random association of opsin alleles in wild groups of red-bellied tamarins (*Saguinus labiatus*) and maintenance of the colour vision polymorphism. *Biology Letters, 1,* 465–468.

EVOLUTIONARY APPROACH: PERCEPTUAL ADAPTATIONS

Anyone with the slightest knowledge of cameras who considers the anatomy of the eye can understand that its purpose is to gather light and focus it on the retina. The eye is the "poster child" for evolution of the human body by natural selection. Yet the implications of evolution for perception are often overlooked. The various senses evolved to serve the survival and reproduction of organisms. This entry discusses how attention to these functions helps us understand perception.

What Is Perception For?

Because philosophy was such an important part of the history of perception, it is often said that the purpose of perception is to gain knowledge. But long before there were philosophers, people had to move about in space without bumping into things, decide what is edible, swallow food without choking, and find their way home. Thus, perception is first of all for action. For example, we are much better at moving about in space and avoiding objects than we are at estimating our distance from those objects.

We could avoid collision with objects by estimating our distance from them and our speed of approach. Instead, we do something much simpler: Because the image on the retina grows at a rapidly increasing rate as we approach an object at a constant speed, we are able to avoid collision by stopping when the rate of expansion reaches a threshold. This simple method is used by pilots landing a helicopter, pelicans diving for fish, and babies blinking at looming objects.

Perception and Fitness

An evolutionary approach leads us to consider the energy requirements of perception. We note that more than half of the human brain is devoted to perception, and the brain consumes about 25% of our metabolic resources. Thus, perception must make a great contribution to our fitness; otherwise, we would spend that energy on some other resource, such as muscles. Humans are properly considered perceptual generalists, in that our senses respond to many kinds of information. Nevertheless, we devote metabolic resources to those sources of information that are most useful to us, and ignore others. Thus, we do not respond to ultraviolet light and are relatively poor at echolocation, which is judging objects and their distances by sounds reflected from those objects. Similarly, sessile animals (those that don't move), such as sponges, lack eyes altogether. A remarkable example of this principle involves sexual dimorphism, where with at least one animal, the insect stylops, the winged male has eyes, but the parasitic female is sightless.

Sensory Ecology

Consideration of the sources of energy available in the world enables us to understand what the senses were designed to do. Vision, for example, responds to electromagnetic energy, which is abundant in the daytime, travels extremely fast in a straight line, and is little affected by air. Still, we use only a small part of the electromagnetic spectrum. Wavelengths longer or shorter than the visible range travel through objects, and so would be less useful. In addition, shorter wavelengths tend to be filtered out by the atmosphere, so there is not much available to be used. Further, shorter wavelengths, such as ultraviolet, X-rays, and gamma rays, tend to be damaging to tissue. So it is not surprising that the eye contains pigments that filter out near-visible ultraviolet rays, which we could otherwise see.

Sound waves travel through air, water, and solids. Because sound travels in the dark, and goes around corners, it is useful at night when there is little light. Further, the relatively slow speed of sound waves makes it possible for us to locate sounds by their time of arrival at the two ears.

Perhaps more than any other sense, audition demonstrates the physical principles that shape the senses. Larger animals such as elephants are capable of producing low frequencies (infrasound), which travel over long distances. By contrast, small insects can only produce high frequencies, which do not travel as far. Whereas the production of sound is limited by physical principles and may be incidental to size of the organism, the reception of sound is necessarily an adaptation; that is, it is the product of selection. The elephant's ability to hear extremely low frequencies (infrasound) permits it to communicate over long distances. By contrast, tiny insects respond to high frequencies that only travel the short distances of significance to them.

The Case of Pain

Pain provides a particularly good example of how evolution selects mechanisms that serve our survival and reproduction, regardless of other consequences. Although we might think life would be better without pain, those rare cases of individuals lacking pain show otherwise. F. C. was a young woman who was born without any pain sensitivity. As a child, she received third-degree burns from kneeling on a hot radiator while looking out the window. She bit her tongue so often that it became deformed, and she eventually bit the end off. But these were not the most serious problem. Her joints deteriorated because she failed to make normal postural adjustments, and this eventually led to a massive infection that took her life at age 29. Fortunately for us, natural selection has given us many sensations that are pleasurable, so that we seek them out. But pleasure and pain both serve our fitness.

Sensory Adaptation Is an Evolutionary Adaptation

The term *adaptation* has two separate meanings that are both highly significant in the present context. An evolutionary adaptation is a mechanism crafted by natural selection to solve problems posed by the environment encountered by ancestral populations during their evolution. So far, this entry has considered the senses as adaptations. One of those mechanisms is sensory adaptation. Sensory adaptation allows us to tune out stimuli that do not provide us with new information needed to cope with the environment. This is the property of adaptation that is generally used to define adaptation in textbooks. But this overlooks that adaptation makes us more sensitive to the changes that do occur.

One of the ways that adaptation makes us more sensitive to changes is by range shifting. You have no doubt had occasion to go into a darkened building such as a movie theatre from bright daylight, and noticed that you had trouble seeing in the reduced light. Then you adapted to the darkness and could see clearly about you. Later when you left the theater, you were temporarily blinded by the light outside. At any one time, our eye responds to only about a two log unit (100 to 1) range of light. By contrast, the range of intensities that our eye is capable of responding to is about 15 log units, or a thousand trillion to 1. Because we are most sensitive to changes in intensity near the level to which we are adapted, range shifting permits us to place our greatest sensitivity to changes where it is most needed. Similar shifts, albeit less dramatic, occur in the other senses. For example, our range of thermal adaptation is only a few degrees because temperatures outside of that narrow range are life threatening.

Other consequences of adaptation include an increase in response to certain other stimuli after adaptation. In the example of thermal adaptation just given, after adaptation to warmth, a previously neutral stimulus will seem cool (as well as pleasant), aiding us in maintaining optimal body temperature. Sensory adaptation should be seen as one of a number of mechanisms that serve the fitness of organisms to the environment.

Modularity of Perception

The senses evolved to deal with specific problems, such as identifying predators and judging where they are looking and where they are going. These particular questions would be expected to lead to the development of a number of specific mechanisms, rather than to a single general-purpose mechanism. Thus, the brain should consist of a number of specialized “organs” corresponding to these tasks, just as we have lungs for breathing and a stomach for digestion. Modularity of perception, then, is to be expected from an evolutionary perspective. Many examples of modularity are seen in perception. The distinction between rod and cone vision is familiar, as are their specializations for different light conditions. Rods are more widely distributed in the retina, thus permitting peripheral vision. They respond more to movement than the cone system, and function in lower light conditions. Cones serve color vision and are more densely arrayed in the center of the retina, permitting better acuity. In addition, two separate processing streams in the cortex serve different sets of functions that have been summarized as “what” (for identifying objects) versus “where” (for locating objects or carrying out actions involving objects). Further, the visual areas of the brain also tend to be segregated into regions that serve special purposes such as perceiving motion, faces, and the like. The auditory structures of the brain are also organized into two separate “what” and “where” systems. Injury or stroke can leave a person with the ability to localize sounds but unable to recognize them, or vice versa.

The many examples of perceptual adaptations indicate that an evolutionary approach to perception leads us to appreciate the organism in its environment and its need to use the available sources of information toward its survival and reproduction.

Donald H. McBurney

See also Animal Frequency and Pitch Perception; Cutaneous Perception; Evolutionary Approach; Modularity; Visual Light- and Dark-Adaptation

Further Readings

Gaulin, S. J. C., & McBurney, D. H. (2004). *Evolutionary psychology* (2nd ed.). Upper Saddle River, NJ: Pearson.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action.* New York: Oxford University Press.

Shepard, R. N. (1992). The perceptual organization of colors: An adaptation to regularities of the terrestrial world? In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 495–532). New York: Oxford University Press.

Warren, W. H. (1995). Self-motion: Visual perception and visual control (pp. 263–365). In W. Epstein & S. Rogers (Eds.), *Handbook of perception and cognition: Perception of space and motion* (2nd ed.). New York: Academic Press.

EXPERIENCE-DEPENDENT PLASTICITY

Experience-dependent plasticity refers to changes in the anatomical or functional organization of the nervous system as a consequence of sensory input or motor output. This can be at the anatomical level of how different neurons project from one area to the next (e.g., how neurons from the eye send their axons to the brain). Alternatively, the gross anatomical projections themselves may not change; rather, the individual synapses are refined, or pruned, so that only restricted recipient neurons are contacted. Finally, there can be instances where functional changes occur in the absence of anatomical changes. In this case, certain synapses are strengthened, whereas other synapses are weakened, with the result that neurons that used to respond to a particular sensory input now respond to a different sensory input (e.g., from touching the index finger to touching the thumb).

Plasticity occurs under a number of different conditions, such as development, neuronal injury, or experience. For example, the amputation of a finger will result in a reorganization of the primary somatosensory cortex so that neurons that used to respond to the amputated finger come to represent the adjacent, intact finger. A second example is following massive injury to a particular sensory structure, such as the eyes leading to blindness or to the cochlea leading to deafness. In these cases, evidence indicates that the parts of the brain that usually represent the missing sensation come to respond to a different sensation, such as visual areas responding to auditory stimuli in blind individuals and auditory areas responding to visual input in deaf individuals.

Plasticity strongly depends on the stimulus history of the individual. This can be the result of experience during development or in adulthood well after all developmental processes have been completed. This entry provides a brief description of both developmental and adult plasticity, providing examples of these plastic changes in different sensory systems.

Normal Organization of Primary Cortical Areas

Experience-dependent plasticity has been described in both cortical and subcortical areas of the nervous system. Although subcortical plasticity is known to occur in many contexts, it remains unclear whether the reorganization of the cerebral cortex is a reflection of the reorganization of subcortical areas, the subcortical reorganization is directed by the cerebral cortex, or some combination of the two. This entry concentrates on the plastic changes in the cerebral cortex and will not address the topic of subcortical plasticity.

In all mammals, distinct cortical areas process the initial inputs to the cerebral cortex for a specific sensory modality. Thus, auditory, visual, and tactile stimuli are processed initially in the primary auditory, visual, and somatosensory cortex, respectively. The bulk of the early work was conducted in these primary areas because they have an orderly and predictable functional organization. For example, there is a "map" of the contralateral body surface in the primary somatosensory cortex, such that neurons that respond to touching the big toe are next to neurons that respond to touching the second toe and neurons that respond to the foot are next to those that response to the ankle. A similar organization is seen in the primary visual cortex, where neurons that respond to one region of space are adjacent to neurons that respond to adjacent regions of space, creating a map of visual space. The auditory cortex is organized somewhat differently, with neurons responding best to a particular tone frequency (pitch) forming a band across the cortex, and neurons that respond best to higher tone frequencies are organized as bands progressively farther from the lower frequency bands. Thus, in the auditory cortex, each frequency is represented by a strip of neurons.

Developmental Experience-Dependent Plasticity

Stimulus-driven activity critically influences the development of sensory systems. If the nervous system does not receive normal sensory input during a defined time in development, termed the critical period, then it will not develop normally. A wealth of knowledge has been gained by experiments exploring the basic principles of experience-dependent plasticity during development, prompted by the classical experiments of David Hubel and Torston Wiesel. The cerebral cortex shows much greater capacity for plasticity during development compared to adult plasticity, which is largely based on exuberance of axonal projections in young animals, which are normally pruned and eliminated in adults, but which allow for a greater range of inputs to the recipient neurons in young animals.

Visual Cortex

If stimulus-driven activity of one eye of a developing mammal is prevented, a condition known as monocular deprivation (MD), there is a dramatic change in the way that the visual cortex develops. Normally, the inputs to the visual cortex form ocular dominance (OD) columns, such that neurons are segregated into those that respond to one eye next to those that respond to the same region of space, but from the other eye. In MD animals, the axons that respond to stimulation of the patched eye are greatly retracted, allowing the axons from the open eye to make more synapses on the cortical neurons. When the patch is removed, the region of the cortex that responds to the eye that was patched is much smaller than is the region of the cortex that responds to the eye that was not patched (Figure 1). This organization persists through adulthood and cannot be reversed, and consequently, that eye is effectively blind, even though it is perfectly fine.

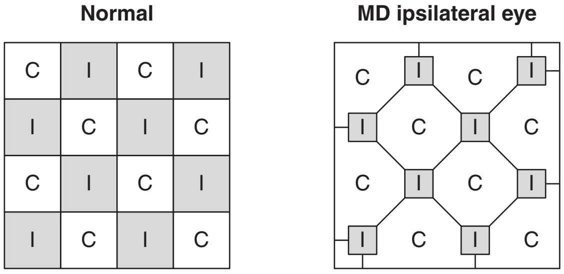


Figure 1 Change in Ocular Dominance Column Organization Following Monocular Deprivation

Note: The deprived eye (in this case, ipsilateral) has a much-reduced cortical territory compared with the intact eye.

Subsequent experiments in the past decades have further defined the critical period where this effect occurs. Similar types of experiments in older or younger cats or monkeys, or if the patching does not occur throughout the critical period, do not result in the same dramatic change. During normal development, experience-dependent activity shapes the anatomical structure of the inputs to the visual cortex as well as the functional connections.

Somatosensory Cortex

A second good example of developmental plasticity is in the representation of the whiskers in the rat somatosensory cortex. Rats are nocturnal animals that depend heavily on inputs from their whiskers to palpate their environment and use them to effectively "see" in the dark. Correspondingly, there is a large representation of the whiskers in the primary somatosensory cortex. Anatomically discrete structures in the cortex where neurons are closely packed are called barrels. A single barrel corresponds to each whisker on the snout on the opposite side, and the neurons in these barrels respond most strongly (although not exclusively) to the whisker corresponding to that barrel. This anatomical and functional structure can be modified during development by removing one or more whiskers. If this is done, the corresponding barrel does not form. If done after the critical period, which is a matter of days after birth, the barrel forms normally anatomically, but the neurons respond to the adjacent whiskers (the corresponding whisker is not there anymore). Thus, the experience of receiving inputs from one whisker establishes the anatomical structure of the barrel. There is plasticity in the responses of barrel neurons in adults, but the anatomical structure remains regardless of the stimulation history and experience.

Auditory Cortex

The auditory system also establishes critical periods during development, with the functional organization of the primary auditory cortex able to change as a function of passive exposure to different sounds during early development. For example, playing noise but not tones to rat pups results in neurons that respond to all frequencies instead of a select few. This must be done at a young age because the mother rat is exposed to the same sounds as the pups, yet the mother rat shows no change in the functional organization of the primary auditory cortex. Again, exposure of certain sounds during the early critical period can result in large-scale changes in the way that the auditory cortex responds.

These results have implications for individuals who are born deaf because there are critical periods of development for the ability to process and understand speech sounds. Speech is arguably the most important sensory stimulus that humans perceive, as without it, the ability to communicate with others is severely impaired. A device called the cochlear implant consists of a set of stimulating electrodes that can be inserted into the cochlea of a deaf individual to simulate some type of hearing. If the device is implanted within the first 4 to 5 years of life of someone born deaf, chances are good that near normal speech understanding will occur. However, if it is implanted later, the device is less effective. In contrast, if a person becomes deaf as an adult, there is no critical period in which the device will work better or worse depending on when it is implanted. In the case of speech perception, the developmental critical period almost certainly influences higher-order cortical areas, well beyond the primary auditory cortex and likely within language areas of the brain such as Wernicke's area.

Adult Experience-Dependent Plasticity

William James, one of the great early psychologists of the late 19th century, noted that adult animals and humans retained the ability to learn new behaviors and would improve on them with practice throughout life. Although this was well known by most people, he was the first to postulate that learning such new behaviors and skills was reflected in changes in the excitability of certain neural structures. This was ahead of his time because the neural structures underlying perception and skill acquisition were still decades away from being identified. The first measurements of such changes at the neuronal level were performed by Sir Charles Sherrington, who noted that electrically stimulating the motor areas of the brains of anesthetized monkeys changed the way that the motor cortex was functionally organized. For example, if he stimulated one location, the monkey's hand would twitch. If he stimulated a few millimeters away, the monkey's elbow would move. If he went back and stimulated the hand place again and again, he found that then stimulating the elbow place would make the monkey's hand move. Thus, by stimulating for several minutes, the elbow part of the motor cortex changed to be incorporated into the hand part. This was one of the first direct demonstrations that just stimulating, or changing the way that the neurons were activated, could cause the adult cerebral cortex to change.

Somatosensory Cortex

The initial modern experiments that investigated experience-dependent plasticity in the cerebral cortex were conducted in the primary somatosensory cortex in the 1980s and early 1990s. These experiments showed that the details of the topographic representation of the body surface, specifically the representation of the fingers in the primary somatosensory cortex of monkeys, could be altered in the adult by a number of different manipulations. For example, most neurons in these areas respond only to touching one finger, and only extremely rarely will they respond to touching the skin of two different (but always adjacent) fingers. However, surgically suturing the ring and middle fingers together for several weeks made neurons that normally responded to only one finger respond to both. This persisted even after the two fingers were again surgically separated, indicating that this was not caused by the sensory receptors growing across the suture line to the other finger. Similar results were shown for the "neurovascular island" transplant. This is a surgical technique to help people who have suffered severe nerve damage in the hand, usually the thumb. The skin, artery, vein, and nerve to a section of skin on the ring finger are transplanted to the damaged part of the thumb. This results in the patient being able to feel the thumb again and a corresponding reorganization of the hand representation in the cerebral cortex. The neurons that used to respond to the ring finger come to respond to stimulation of the thumb (the same skin); however, these neurons are adjacent to those that respond to the index finger, not the pinky as before.

Subsequent experiments verified that these experience-dependent forms of plasticity had functional consequences. These included a study where monkeys were trained to perform a frequency discrimination task to vibrations on a single finger. Over several weeks, the monkeys improved their performance, and the area in the cortex where neurons responded to touching the skin that was stimulated grew larger. In addition, how well the neurons responded to the particular frequencies was sharpened and was directly correlated with how well that particular monkey could perform the task. Thus, these experiments showed that repeating a tactile discrimination task (practice) resulted in both an improvement in performance over time and an expected change in the functional response properties of cortical neurons. These findings have since been extended over a large range of different preparations. For example, human imaging experiments have shown that people who play musical stringed instruments have a larger representation of the left hand in the primary somatosensory cortex compared with non-musicians.

Auditory Cortex

A similar result is seen in the auditory system. Monkeys that learn to discriminate between tones at particular frequencies have an expansion of neurons responding to those frequencies in their primary auditory cortex. In this case, the expansion of the trained frequencies was directly correlated with the ability of the individual monkey to perform the frequency discrimination. An important finding from these training experiments in the tactile and auditory cortex is that attention was required for the plastic changes to occur. The monkeys trained to perform the tactile discrimination task heard in the background the same tone frequencies as the other monkeys did, but had completely normal auditory cortical representations (but big changes in their somatosensory cortex). Similarly, the monkeys performing the auditory frequency discrimination task were also stimulated on their fingers exactly the same way as were the monkeys trained on the tactile discrimination task, but again it was irrelevant. These monkeys had completely normal representations of their hands in the somatosensory cortex (but big changes in their auditory cortex). Thus, passive stimulation in these contexts does not drive plastic changes in the cerebral cortex.

An experimental technique that results in cortical plasticity in the absence of overt attention is electrically stimulating a region in the brain that is known to release the neuromodulator acetylcholine throughout the cerebral cortex. Pairing stimulation of this structure with passive exposure to a single tone frequency results in an increase in the representation of that frequency, similar to that seen with behavioral training. Similarly, pairing stimulation with specific temporal sequences will result in an enhancement of the temporal processing ability of cortical neurons. These studies indicate that neuromodulators have a distinct role in the generation of cortical plasticity, with acetylcholine used in this example, but other neuromodulators such as norepinephrine and dopamine have also been implicated in other aspects of cortical plasticity.

Visual Cortex

Training-induced cortical plasticity is seen in the visual cortex as well and has the advantage that the parameters that are altered can give insights into the level that the plasticity is effective. For example, perceptual gains that are restricted to a single eye must be manifest in monocular regions of the nervous system, and not in binocular regions (areas beyond the primary visual cortex). Such training experiments have shown that plastic changes can occur as early as the primary visual cortex (V1), and these changes can be an expansion of the representation of the relevant stimulus features, as in the auditory cortex. Thus, across somatosensory, auditory, and visual cortical areas, the functional organization can be altered as a consequence of training at a particular sensory discrimination task. These changes are correlated with the ability to perform the particular task, and thus in the ability to perceive the stimuli, and it is likely that this experience-dependent plasticity is the underlying neuronal mechanism of skill acquisition with practice.

Gregg H. Recanzone

See also Auditory Processing: Central; Cochlear Implants: Technology; Perceptual Learning; Visual System Structure

Further Readings

Fox, K., & Wong, R. O. L. (2005). A comparison of experience-dependent plasticity in the visual and somatosensory systems. *Neuron, 48*(3), 465–477.

Karmarkar, U. R., & Dan, Y. (2006). Experience-dependent plasticity in adult visual cortex. *Neuron, 52,* 577–585.

Moucha, R., & Kilgard, M. P. (2006). Cortical plasticity and rehabilitation. *Progress in Brain Research, 157,* 111–122.

Pantev, C., Ross, B., Fujioka, T., Trainor, L. J., Schulte, M., & Schultz, M. (2003). Music and learning-induced plasticity. *Annals of the New York Academy of Sciences, 999,* 438–450.

EXTRASENSORY PERCEPTION

Extrasensory perception (ESP) is the most common term for a number of related claims, such as clairvoyance (knowledge of events other than by the senses), telekinesis (affecting events other than by the body), and precognition (knowledge of events before they happen). Although such claims go back many centuries, the modern period began with the founding of the (British) Society for Psychical Research in 1882. The founders were prominent intellectuals who were motivated to find evidence for phenomena that materialism could not explain, including life after death. Thus, much of the work in this area has a quasi-religious motivation. This entry describes the phenomena attributed to extrasensory perception and the controversy surrounding these phenomena.

The history of this area is one of repeated claims based on some methodology, followed by failures to replicate. Among the prominent workers have been J. B. and Louisa Rhine, who conducted laboratory work with card guessing in the United States during the 1940s. In England, S. G. Soal performed similar studies about the same time. The methodological weaknesses in this research led to much controversy.

More recent research has included the Ganzfeld situation, in which an isolated individual attempts to send information about a target to the subject in a type of sensory isolation effected by putting half a Ping-Pong ball over each eye. Also, Robert Jahn conducted many studies in which subjects attempted to affect the output of random number generators. These two lines of work have produced large data sets that have been analyzed using meta-analysis, a method of considering many separate studies together. One problem with such research comes from what is called the law of truly large numbers. As Percy Diaconis and Fred Mosteller point out, when a sample becomes extremely large, extremely rare events tend to happen. In the present case, tiny effects caused by design flaws that would be inconsequential in ordinary research could be responsible for statistically significant results when the data set becomes extremely large. In the usual sorts of research, researchers study larger effects so that experimental design and methodology are not so critical.

These studies are only a small part of a large number that have been published over many years. The literature as a whole provides inconsistent evidence of paranormal effects, indicating the existence of a file-drawer problem: Investigators conduct a great many studies but publish mostly those that give positive results.

The judgment of the scientific community concerning the validity of the field of ESP may be seen in the 1988 report of the U.S. National Academy of Sciences that found no justification for the belief in parapsychological phenomena based on 130 years of research.

It is significant that the research in ESP has been conducted largely by people with no particular expertise in sensory processes and perception, the areas of research most closely related to the field. Thus, the typical studies have various flaws that would be avoided by those with relevant training. For example, researchers Russell Targ and Hal Puthoff employed a so-called soundproof room to isolate their subject, Uri Geller, who was famous for bending spoons and reproducing figures sketched by others. Targ and Puthoff were apparently unaware that the sound attenuating properties of such rooms are far from total and that it would be a simple matter to communicate by ordinary means through the walls. In fact, Geller's assistant was outside the room during the session, and evidently communicated to Geller what he should draw.

A particularly revealing example is given by the work on "digital color sensing" in the 1960s. It was claimed that people could detect the color of items in the dark using only their skin. Walter Makous, a visual scientist, recognized that the phenomenon, if true and if based on color, per se, would require unknown sensory capacities. Others had claimed that failures of the blindfold system must be responsible. But Makous was able to show that the temperature sensitivity of the skin was sufficient to detect differences in the heat that was reflected back to the skin by surfaces of different color. First, he showed that this was theoretically possible given the physics of the situation and the thermal sensitivity of the skin. Specifically, the longer wavelengths (the so-called warm colors, especially red) reflect more infrared energy, and thus feel warm. Shorter wavelengths (the so-called cool colors, especially blue) reflect less and thus feel cool. Then he did an experiment in which subjects discriminated between a front surface mirror and a black surface placed near the skin. These conditions maximize the differences possible when relying on the thermal sensitivity of the skin as the signal. The results were consistent with his calculations. His paper put a sudden and virtually complete end to this line of work. The paranormal community had little interest in understanding how the senses operate, only in the possibility of finding a phenomenon that could not be understood by ordinary science.

Researchers have questioned the findings of parapsychological research, both on the grounds of methodological problems and the failure to replicate, and philosopher C. D. Broad has noted that the proposed effects would violate the limiting principles of science. These principles are seldom discussed but nevertheless form the basis of all science: (a) that effects cannot precede causes, (b) the mind depends on the brain, (c) the world affects the mind only via the body, and (d) the mind affects the world only via the body. These are contradicted respectively by claims for (a) precognition, (b) seeing objects during out of body experiences that would be otherwise hidden from view, (c) extrasensory perception, and (d) telekinesis.

Although science is a highly interconnected enterprise in its institutional structure, methods, results, and theoretical considerations, the field of ESP stands apart. For example, ability to communicate by ESP has not been shown to be affected by distance, whereas ordinary physical signals decrease in intensity over distance, following the inverse square law. Nor is there any consistent relation between ESP and emotion. On the hypothesis that ESP would be more likely under great stress, some have looked for evidence of naturally occurring paranormal communication in wartime or natural disasters. Others have taken the opposite tack, supposing it would be better during relaxation, as in the Ganzfeld studies. Unlike normal science, the field of ESP shows a complete lack of progress. A discussion of the most recent studies reveals no newer principles or understanding than we had 50 or 100 years ago.

Given that positive results have been so hard to find, it may not be surprising that some of the apparently strongest evidence has been the result of fraud. George Price published two articles in which he proposed that many of the most famous results must be fraudulent and suggested how they may have been produced. Although the paper was assailed as speculation, and Price later even apologized for making unfounded accusations, further events showed his article to be remarkably prescient. For example, the work of Soal, in particular, was much later shown to have been the result of his altering the data. Walter J. Levy, Rhine's chosen successor in his lab, was caught altering data in a supposedly automated study of the ability of rats to predict pleasurable brain stimulation.

Donald H. McBurney

Further Readings

Alcock, J. E. (1985). Parapsychology as a "spiritual" science. In P. Kurtz (Ed.), *A skeptic's handbook of parapsychology* (pp. 537–565). Buffalo, NY: Prometheus Books.

Bem, D. J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin, 115,* 4–18.

Broad, C. D. (1949). The relevance of psychical research to philosophy. *Philosophy, 24,* 291–309.

Diaconis, P., & Mosteller, F. (1989). Methods for studying coincidences. *Journal of the American Statistical Association, 84,* 853–861.

Druckman, D., & Swets, J. A. (Eds.). (1988). *Enhancing human performance: Issues, theories and techniques.* Washington, DC: National Academy Press.

Hansel, C. E. M. (1980). *ESP and parapsychology: A critical re-evaluation.* Buffalo, NY: Prometheus Books.

Hyman, R. (1985). The Ganzfeld Psi experiment: A critical appraisal. *Journal of Parapsychology, 49,* 3–49.

Jahn, R. G. (1995). "Out of this aboriginal sensible muchness": Consciousness, information and human health. *Journal of the American Society for Psychical Research, 89,* 301–312.

Makous, W. L. (1966). Cutaneous color sensitivity: Explanation and demonstration. *Psychological Review, 73,* 280–294.

Price, G. R. (1955). Science and the supernatural. *Science, 122,* 359–367.

Price, G. R. (1956). Where is the definitive experiment? *Science, 123,* 17–18.

Price, G. R. (1972). Apology to Rhine and Soal. *Science, 175,* 359.

EYE: STRUCTURE AND OPTICS

The human eye is a simple, but extremely robust, optical instrument. It produces images of the visual world onto the retina with a quality that matches or exceeds the resolution of the neural circuitry. The eye is composed of only two positive lenses, the cornea and the crystalline lens, that produce real images of the world on the retina, initiating the visual process. Traditionally, the eye has been regarded as an equivalent of a camera, with the ocular optics acting as the lens and the retina as the film (or the sensor). Compared with artificial optical systems, often formed by many lenses to improve image quality, the eye is simple, but well adapted to the requirements of the visual system. Lenses form images of objects by bending the light. The refractive index is a measure of how much a lens material changes the direction of the light, from a value of 1 for air to a typical figure of 1.5 for glass. The eye has to form high-definition images of a large field of view for objects placed at different distances using transparent living tissue, instead of glass. Because of the nature of light and its own characteristics, the eye imposes the first physical limit to vision. Vision will be good if the images formed on the retina have sufficient quality. However, if the retinal images are blurred or fuzzy, the visual system will not function properly and vision will be poor. This entry describes the structure and optical properties of the eye.

Structure of the Eye

The eye in adult humans is approximately a sphere about 24 millimeters (mm) in diameter. Externally, it is covered by a resistant and flexible tissue called sclera, except in the anterior part where it turns to the transparent cornea that allows the light to pass into the eye. Within the sclera are two other layers: the choroid, which is a mesh of blood vessels to provide nutrients, and the retina, where the light is absorbed by the photoreceptors as the second step of the visual process. The eye can be moved by the action of six external muscles that permit fixation and the scanning of the visual environment.

The light reaching the eye is first transmitted and refracted by the cornea. This is a thin transparent layer free of blood vessels of about 12 mm in diameter with slightly more than 0.6 mm thickness in the central part. An aqueous tear film on the cornea ensures that the first optical surface is smooth to provide the best image quality. After the cornea, there is a small space, called the anterior chamber, that is filled with a water-like substance, the aqueous humor. In the posterior area of this space is the iris, a muscle with a central hole the size of which depends on the iris's contraction. The iris is a diaphragm with characteristic color, depending on the amount and distribution of pigments. The pupil is the variable opening in the center of the iris that limits the amount of light passing into the eye. The pupil size changes with the ambient light, from less than 2 mm in diameter in bright light to more than 8 mm in the dark. Other factors, such as emotions or certain drugs, also affect the size of the pupil. The pupil controls retinal illumination and limits the rays entering the eye to improve the images.

After the iris is the crystalline lens, which combines with the cornea to produce the images on the retina. To ensure transparency, similar to the cornea, the crystalline lens has no blood vessels, with the surrounding aqueous humor providing nutrients. With normal ageing, crystalline lenses tend to become less transparent, which diffuses light and degrades the retinal image. In advanced stages of this situation, a condition called cataracts, the lens is nearly opaque, leading to a low quality of vision. The standard, and successful, surgical procedure consists of replacing the cataractous lens with an artificial implant, called an intraocular lens, that restores the eye's ability to form clear images on the retina.

The cornea is a lens with fixed optical power, and the crystalline lens is an active optical element. It changes its shape, which modifies its optical power, and therefore the whole eye's power. This is the basis of the mechanism of accommodation that allows the eye to focus in the retina objects placed at different distances, from distant to near. The lens is suspended in a bag with a set of muscles attached, the ciliary muscles. The action of these muscles permits the lens to increase or decrease power. The ability to keep near objects in focus declines continuously with age. After 50 years old, the crystalline lens cannot change its shape anymore because of the modifications of its mechanical properties. This condition is called presbyopia and can be alleviated by using different optical tools, such as bifocal, progressive, or contact lenses.

After the light is finally refracted by the lens, it enters the area called the posterior chamber, which is filled with a transparent and viscous fluid, the vitreous humor. After passing through this area, the light reaches the retina, the thin layer of neural tissue covering the back of the eyeball that acts as a screen where images are formed and the light is converted in electrochemical signals. The human retina has a central area, the fovea, where photoreceptors are densely packed to provide the highest resolution. This area is used for demanding visual tasks, such as reading. The eyes continuously move to fixate the desired details into the fovea. The peripheral parts of the retina render lower resolution but are important for movement detection and to first locate objects in the visual field. The eye's optics produces retinal images that match the foveal requirements in spatial details but have better resolution than that required in the peripheral retina.

Optical Properties of the Eye

The eye as an optical instrument is composed of two lenses, the cornea and the crystalline lens, the pupil to limit the light entering the eye, and the retinal plane, which acts as a screen. The cornea is approximately a spherical section with a refractive index of 1.377. The largest difference in refractive index occurs from the air to the cornea (actually the tear film), so this accounts for most of the refractive power of the eye, over 70% on average. The rest of the optical power to bring the images into focus at the retina is provided by the lens. This is a biconvex lens with the anterior surface (the one closer to the cornea) and the posterior surface with radii of 10.2 mm and –6 mm respectively. The internal structure of the lens is layered (similar to an onion), which produces a nonhomogeneous refractive index, larger in the center than in the periphery, and with a mean value of 1.42.

An average eye with a distance from the cornea to the retina (called axial length) of 24.2 mm will image distant objects precisely in focus into the retina (see Figure 1). This ideal eye is called an emmetrope and has an optical power of about 60 diopters (a diopter is a unit of measurement of optical power given by the inverse of the focal length in meters). In a young emmetropic subject, near objects will be also in good focus because of accommodation. However, most eyes have neither the adequate optical properties nor the dimensions to achieve perfect focus; they are affected by optical refractive errors. In these cases, the images formed in the retina are blurred, typically imposing a lower limit to visual perception. Refractive errors are classified as myopia (or near-sightedness) when the images of distant objects are focused in front of the retina; hypermetropia (or far-sightedness) when the distant objects are focused behind the retina. Typically, myopic eyes are longer and hyperopic eyes are shorter than are those of an emmetropic eye. A visual deficit has an optical origin if vision is restored after adequate optical correction. Refractive errors are corrected by placing either negative or positive lenses in front of the eye, as eyeglasses or contact lenses. Another permanent alternative is re-shaping the cornea by laser refractive surgery. In every person, the eye is not rotationally symmetric as ideally described here. The optical surfaces are not exactly spherical in shape, and they are not perfectly aligned. A common manifestation of these facts is the presence of ocular astigmatism: the retinal images of a point source are two perpendicular lines. If astigmatism is severe, it can degrade image quality, compromising vision. The correction of astigmatism requires the use of cylindrical lenses or refractive surgery techniques.

Even eyes free of any refractive error, without defocus and astigmatism, do not produce perfect images. The retinal image of a point source is not another point, but an extended distribution of light (see Figure 1). This sets the minimum detail that can be imaged in the retina and then visual resolution. Details subtending less than 1 minute of visual angle (a letter detail of 0.1 mm for a reading distance 40 cm) are readily resolved by the eye and well discriminated in the fovea. This corresponds to the standard value of 20/20 (or 1 in decimal notation) visual acuity.

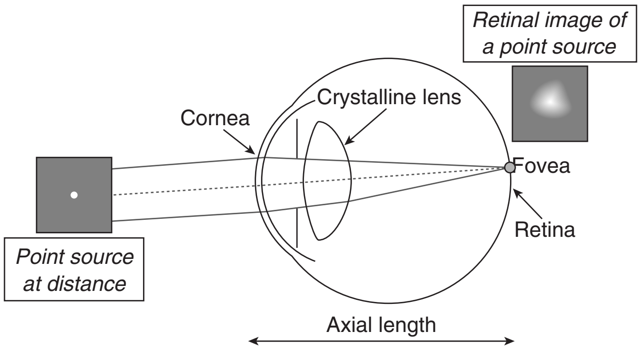


Figure 1 Diagram of the Optics of the Human Eye With Description of Main Components

Several factors are responsible for the degradation of the retinal images: diffraction of the light in the eye's pupil, optical aberration, and intraocular scattering. Diffraction blurs the images formed through instruments with limited apertures, such as the pupil, because of the wave nature of the light. The effect of diffraction in the eye is small and only noticeable with small pupils. It can be perceived as halos when seeing distant lights through a small pinhole aperture. Aberrations affect all optical instruments, particularly the eye. Light rays entering the eye at different parts of the pupil are not focused at the same retinal locations, causing an additional blur in the image. Two of the more important aberrations in the eye are spherical aberration and coma. An optical system is affected by the coma aberration when it forms images of a point source with a comet-like shape. The impact of ocular aberrations in image quality is more significant for larger pupil diameters. The aperture of the eye ranges from f/8 to f/2, which can be compared with values in a camera. The amount of aberration for a normal eye with about an f/4 aperture is approximately equivalent to less than 0.25 diopters of defocus, a small error typically not corrected when dealing with refractive errors.

In the normal young eye, the optical properties of the two ocular components are somehow tuned to produce an improved overall image quality. In particular, coma and spherical aberration are partially corrected in the eye, similar to an aplanatic artificial optical system. It is rather extraordinary to realize how the optics of the eye is optimized with the crystalline lens correcting the corneal spherical aberration and with the proper choice of shapes to reduce coma. During normal aging, this balance is partially disrupted leading to an increase of the eye's aberrations and a degradation of the retinal image quality.

An additional source of degradation of the retinal image is caused by the spectral content of the white light and the dispersive nature of the eye media. The optical power of the eye depends on the wavelength (color) of the light, so an eye perfectly focusing green light in the retina will be myopic for blue and hyperopic for red light. This is known as chromatic aberration and causes a colored blur in the retinal images, although with a limited impact on vision.

Despite all those optical imperfections, the eye represents an optimized design solution rendering a stable retinal image quality for a variety of geometrical configurations and meets the exquisite demands of the visual system.

Pablo Artal

See also Eyes: Evolution of; Retinal Anatomy; Visual Acuity

Further Readings

Artal, P., Benito, A., & Tabernero, J. (2006). The human eye is an example of robust optical design. *Journal of Vision, 6,* 1–7.

Artal, P., Guirao, A., Berrio, E., & Williams, D. R. (2001). Compensation of corneal aberrations by internal optics in the human eye. *Journal of Vision, 1,* 1–8.

Atchison, D., & Smith, G. (2000). *Optics of the human eye.* Boston: Butterworth-Heinemann.

Le Grand, Y., & El Hage, S. G. (1980). *Physiological optics* (Springer series in optical sciences). Berlin: Springer.

Pedrotti, F. L., Pedrotti, L. S., & Pedrotti, L. M. (2007). *Introduction to optics* (3rd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

EYE AND LIMB TRACKING

Most daily activities—whether at work, play, or home—involve eye and limb movements. Sometimes the association between the two is direct—for example, when we look at the handle of a coffee cup to reach out to grasp it. Sometimes the association is indirect, as in driving, where one looks at objects on the road while using the hands and feet to control the car. Probably the most studied aspect of this association is eye-hand coordination. As with most such topics, outside of the laboratory we usually take eye-hand coordination for granted unless we remark on the "good eye-hand coordination" of an athlete, or notice a problem that arises in an individual with neurological disease or damage. But how does eye-hand coordination work? Eye-hand coordination has both temporal aspects and spatial aspects. The brain needs to properly control the order and timing of eye and hand movements, and it needs to account for changes in eye position to use spatial vision to guide the hand. This entry reviews the basic aspects of eye-hand coordination and some associated clinical problems.

Behavioral Aspects

When we fix our vision on some point in space, the direction of the line from our eyes to that fixation point is known as our *gaze direction.* Our sense of sight provides highly detailed information about objects along this gaze direction. Therefore, when exploring novel environments, quickly moving, unexpected, or otherwise salient objects that we might want to evaluate in greater detail will tend to attract our gaze. However, when vision is used to guide the limbs for planned action, the nature of eye movement tends to be more structured. In this latter case, eye movement appears to be optimized to obtain information that allows for the most efficient movement of our limbs. For example, if we are making a peanut butter sandwich, we might direct our eyes to a knife so we can determine its location precisely enough to pick it up. Once the location of a given target (e.g., the knife) has been gathered, the eyes initiate a saccadic movement to the next target (e.g., peanut butter jar) at about 100 to 200 milliseconds (ms) before the limb actually reaches the first target. This is likely because visual information from this time period cannot be processed quickly enough to be of use in guiding a limb to the first target, but planning for the next target *can* begin. In some situations, however, gaze direction remains locked to a target until the instant the limb reaches it. This is most easily seen in pointing, where the eyes fixate on the pointing target until the pointing motion is complete. In addition to target location acquisition, our eyes are used to monitor complicated motor tasks and search for errors of movement that can be corrected. Thus, we may watch what our hands are doing as we attempt to spread peanut butter on bread or cut the resulting sandwich into equal halves.

These qualitative observations have been measured and quantified experimentally. For example, when subjects are asked to perform simple tasks that involve grasping and moving of single objects, it is found that their eyes tend to fixate only on locations that will become the targets of limb motion, or on objects that could become obstacles to such motion. When such tasks involve direct reaching to a target, it is found that gaze usually begins shifting to the target at about 100 to 200 ms after the target becomes visible, and the hand begins movement about 100 ms later. Typically, a subject's gaze will reach a visual target at a few hundred milliseconds (usually less than 500 ms) before the hand arrives. This is thought to be enough time for visual feedback (discussed further later in this entry) to aid in accurate reaching.

Although people tend to direct their gaze to reaching targets well before the hand comes near the target, this pattern sometimes occurs in reverse. In 'naturalistic' experiments, subjects occasionally will begin reaching for a target up to a full second before gaze is directed to it. Detailed visual information about a target can only be obtained when it is near where the observer is looking, so this begs the question of how the brain knows where to direct the reaching movement. One possibility is that target location information has been remembered from previous visual fixations. Two ways in which such remembered information could be represented in the brain are discussed in the next section.

The eye is not always used to guide limb motion. Sometimes it works the other way around. For example, when we look for our keys in near darkness, we might find them first with our hands by groping around on a table. To be sure that we have found our car keys instead of our house keys, we might then bring the keys toward our eyes and direct our gaze at them. In this case, gaze is directed based on knowledge of where our limbs are in personal space. Empirical experiments have shown that in such cases gaze shifts are fairly accurate, except for certain errors that appear to arise from a lack of subconscious knowledge of the complexity of the limb system.

Visuomotor Transformation

Any visual information that reaches our brain must do so through our eyes. If such information is to be used to generate behavior, then it must also pass through a transformation process. For example, if someone sees a button in his or her field of view, then the light reflecting off that object will fall on a small region of the person's retina. Before the person can press the button, the retinal location of its image must be converted into a temporal sequence of muscular contraction intensities that will bring the fingertip to it. At this point, we are discussing a purely visual *feedforward* process in which knowledge of the initial conditions (the target location, eye position, etc.) is used to determine all subsequent movements that will be made to reach it. The advantage of this mechanism is that it can be fast, but it must be carefully calibrated and it might make mistakes if those initial conditions have changed. Alternatively, the distance and direction between the retinal image of the effector (fingertip) and the retinal image of the object of interest (button), known as the retinal error, could be sampled by the brain at various times throughout the movement sequence. By using this dynamic information, the brain could provide online corrections to the effector movement. Such a process is known as visual *feedback* because the effector output at one instant in time becomes part of the brain's input at the next instant. The advantage of this mechanism is that it can be accurate, but it is relatively slow because the delay of sensory feedback in the brain is significant.

Although the precise extent to which visual feedforward and feedback processes are used to guide movement to a visual target is not fully understood, both processes are involved. In the case of feedforward, people placed in complete darkness and unable to see any of their own limbs can still point to a small, illuminated target with reasonable accuracy. Because no visual feedback information is available in this case, only visual feedforward and *nonvisual* feedback mechanisms can be acting. However, these other feedback mechanisms are known to be relatively slow compared with how quickly people can point, indicating that purely visual feedforward mechanisms can produce accurate pointing. In the case of visual feedback, it is known that removing visual information about the position of the limbs during reaching movements does reduce the accuracy of those movements, thus demonstrating the existence of these feedback mechanisms.

That visual feedforward mechanisms can be used in isolation to produce rapid and accurate reaching movements tells us that the brain must have an accurate internal knowledge about how the signals it sends to the effector muscles will affect their subsequent position. Computer modeling work has shown that the brain seems to account for the body's detailed geometry, including how the eyes are oriented in the head and how the head is oriented with respect to the body when transforming visual position information from the retina into arm movements (see Figure 1a).

The transformation from retinal position to arm movement is only one type of visual transformation made by the brain. As mentioned in the previous section, people can reach accurately to targets that are not currently fixated, or that are not even currently visible! That they can do so implies that the brain must be relying on a memory of the target location from previous fixations. It has been argued that before the brain can maintain a memory of a target location after a subsequent gaze shift, the brain must transform this location information into a stable frame of reference. For example, a person might remember where an object from a previous fixation is relative to his or her head. This head-centered information will not change with eye movement, so accurate reaches can be made based on it. However, evidence suggests that the brain does something less intuitive in at least some circumstances. In this alternate framework, the target location is remembered relative to the eye. During a shift of gaze, this location information is updated as the eye moves so that the brain "knows" where the reach target is relative to the eye at all times. Only when a limb movement to the target is required does this eye-centered representation of location go through the feedforward transformation process required to generate limb movement. This gaze-centered updating scheme is probably not the only mechanism for spatial memory and action—for example, the brain also stores the relative locations of objects and makes use of body-centered coordinates for some tasks—but it likely reflects a default mechanism for simple visuomotor tasks that provides a framework for other mechanisms.

Cortical Mechanisms

The path followed by visual information through the brain is complex, and the information itself is broken down and distributed across numerous neural circuits. To begin, signals from the retina flow along the optic nerve and undergo preprocessing in various subcortical structures before reaching the

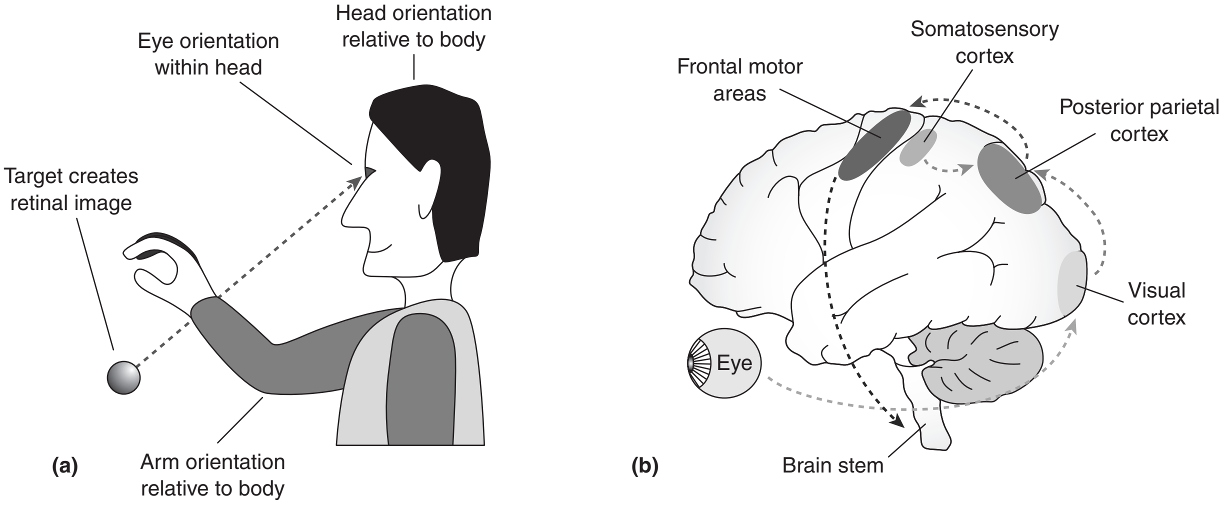


Figure 1 Limb-Tracking Information and the Brain

Source: Adapted from Blohm, G., & Crawford, J. D. (2007). Computations for geometrically accurate visually guided reaching in 3-D space. *Journal of Vision, 7*(5), 1–22.

Notes: (a) Some of the variables required by the brain to transform target location into a reaching plan. (b) Simplified flow of information in the feedforward transformation process. Visual information from the retina flows to the primary visual cortex, then to the parietal cortex where it is combined with information about the body and its current configuration from the somatosensory cortex. This information can then be used by frontal motor areas to generate movement of the appropriate effectors.

primary visual cortex in the occipital lobe (see Figure 1b). Leaving the occipital lobe (visual cortex), visual signals that flow to the parietal lobe are associated with direct action relating to environmental entities. In this area of the cortex, visual information is further separated and transformed into representations of the locations of possible targets for eye and effector (e.g., hand) movements. This is discussed further later in this entry. Information from the parietal cortex is then used by structures in the frontal cortex to generate movement plans that are finally conveyed to the brain stem structures that control the appropriate muscles.

Focusing on the parietal cortex, an area known as the parietal reach region (PRR), which encodes the locations of future reaching targets relative to the eye, has been identified in the monkey. Neuroimaging studies have identified a possible PRR analogue in the human parietal cortex as well. These studies indicate that the putative human PRR also encodes reach target locations in an eye-centered frame and that such location representations are continuously updated as the eyes move. Thus, the PRR might serve as the neural substrate for the eye-centered updating. Another area of the parietal cortex, known as the lateral intraparietal (LIP) area in monkeys and the parietal eye fields (PEF) in humans, encodes the locations of targets for future eye movements. As with the PRR, the PEF seems to maintain an eye-relative representation of target location, even when the target is lost from view and eye movements are made. In addition to representing target locations, neurons in monkey PRR and LIP respond to planned movements when no target location information is yet available. In particular, greater numbers of PRR neurons tend to respond when reach movements are planned, and greater numbers of LIP neurons respond when eye movements are planned, but both areas do show responses to both types of planned movement.

The eye-relative representations of the PRR and LIP/PEF may seem counter-intuitive given that the brain needs to know how to drive the muscles for reaching targets, but it has been argued by some that information about locations of visual objects stays in eye-relative form until it is needed for action. At that point, it must be converted into a form that is more closely associated with the effector limb. Neurons in the parietal cortex of monkeys have been found to represent a mixture of where a target's image falls on the retina and how the eyes are oriented in the head. More specifically, a given neuron will exhibit an elevated firing rate only when the image of a target falls on a small, isolated region of the retina specific to that neuron, with the intensity of this firing response being modulated by the position of the eye in the head. Neurons with this latter property are said to exhibit eye-position-modulated gain fields. Neurons with other gain field properties have also been found in the parietal cortex. These findings are particularly important because computer models show that such gain field neural responses could form the basis of the feedforward transformation between eye-relative representations of location and the kinds of body-centered signals needed to control the limbs. Moreover, both single-unit recordings in animals and brain imaging studies in humans have demonstrated that activity in these parietal regions are remapped during eye movements as predicted by the gaze-centered updating model of spatial memory.

Clinical Aspects

Eye-hand coordination involves vision, eye movements, and arm movements—which in turn involves much of the brain—so it can be negatively affected, either directly or indirectly, by many types of disease process or injury. In the simplest case, the eyes, optic nerves, or primary visual cortex could become damaged, resulting in blindness in all or part of an individual's normal field of view. Similarly, damage to the limbs, spinal cord, brain stem, or motor cortex could result in an inability to use some or all of the limbs. However, deficits in eye-limb coordination are not always so extensive. For example, strabismus is one of the most common visual disorders that children are born with. In this usually congenital condition, the eyes do not align properly so that the affected individual can look directly at an object with one eye only, and the other eye is directed at a slightly different location in space. When a person with normal vision looks at an object, the image made on the retina of one eye is in a slightly different location than it is on the retina of the other eye because of the physical separation between the two eyes. This *binocular disparity* is an important source of information that the brain uses to determine how far away nearby objects are. Depending on the severity of misalignment, strabismus usually results in a lack of ability to use this depth cue and thus reduces the accuracy of reaching movements in depth.

Many other disorders of eye-limb coordination are the result of damage to visual and motor areas of the brain. Take, for example, damage to the parietal cortices, which gives rise to deficits in eye-hand coordination that we can now understand given the physiological knowledge described earlier. Such damage can arise from strokes, traumatic brain injury, tumors, multiple sclerosis, and so on, but the nature and severity of the resulting deficits generally depends on the sizes and locations of lesions rather than on their ultimate cause. An easily understood example of such a deficit is known as optic ataxia. In this condition, patients with damage to certain regions of the parietal cortex have difficulty in reaching directly to visible targets, even though they can visually identify such objects with ease. The degree of impairment varies with lesion size and location, and is generally higher when objects are in the patient's peripheral vision. Optic ataxia sometimes occurs as part of a trio of deficits, known as Bálint's syndrome, that also includes ocular apraxia and simultagnosia. Patients with ocular apraxia have difficulty directing their gaze to objects of interest, whereas those with simultagnosia are thought to be able to perceive only one visual object at a time. It has recently been shown that some of the errors related to optic ataxia can be explained in terms of the gaze-centered spatial updating signals that have been observed in the parietal cortex. Thus, after the patient views an object, an eye movement can either improve or degrade his or her reaching performance, depending whether it shifts the internal image of the object into the good or the damaged part of the parietal cortex.

Because motor output is a complex process that relies on widespread neural circuits, damage of widely varying etiology to any number of brain regions can interfere with this process. Furthermore, although various behavioral aspects of eye-limb coordination have been studied, and some progress has been made in understanding how the brain maintains and transforms visual information into appropriate movements, much remains to be learned. For example, the gaze-centered updating model has allowed us to understand several aspects of how stored information from the eyes and hand remains aligned, but we do not yet know how the brain integrates these simple visuomotor mechanisms with more complex perceptual and cognitive processes.

John Douglas Crawford and Patrick Byrne

See also Action and Vision; Perceptual-Motor Integration; Reaching and Grasping

Further Readings

Andersen, R. A., & Buneo, C. A. (2002). Intentional maps in posterior parietal cortex. *Annual Review of Neuroscience, 25,* 189–220.

Ballard, D. H., Hayhoe, M. M., Salgian, G., & Shinoda, H. (2000). Spatio-temporal organization of behavior. *Spatial Vision, 13*(2–3), 321–333.

Blohm, G., & Crawford, J. D. (2007). Computations for geometrically accurate visually guided reaching in 3D space. *Journal of Vision, 7*(5), 1–22.

Crawford, J. D., Medendorp, W. P., & Marotta, J. J. (2004). Spatial transformations for eye-hand coordination. *Journal of Neurophysiology, 92*(1), 10–19.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*(1), 20–25.

Khan, A. Z., Pisella, L., Vighetto, A., Cotton, F., Luauté, J., Boisson, D., et al. (2005). Optic ataxia errors depend on remapped, not viewed, target location. *Nature Neuroscience, 8*(4), 418–420.

EYE MOVEMENTS: BEHAVIORAL

Movements of the eyes are crucial for being able to see a clear and sharp image of the world. Understanding why this is the case requires considering some facts about vision. An optical image of the visual environment is formed on the retina of each eye, much like the image in a digital photographic camera. Because of properties of both the retinal receptors and the brain, the quality of human vision is inhomogeneous across space. Visual resolution is best in a small, central portion of the retina (the *fovea*), and falls steeply toward the periphery. Movements of the eyes are used to bring images of interesting objects to the fovea, where they can be seen with optimal clarity. Movements of the eye are also responsible for ensuring that the selected images do not leave the fovea, or move across the retina too rapidly, each time the head moves.

These functions are carried out by three main types of eye movements:

1. *Compensatory eye movements.* Humans are always moving their heads. The head is never completely stationary and moves to some extent even when we try to hold as still as possible. Compensatory eye movements, approximately equal in size and opposite in direction to the head movements, are generated continuously to counteract these head movements and keep the image relatively stationary on the retina. Compensatory eye movements are largely automatic reflexes, controlled by evolutionary old systems. These movements are functionally comparable to the image stabilization systems, built into many modern cameras, that enable sharp snapshots despite unsteady hands.

2. *Saccadic eye movements.* Saccades are rapid, stepwise rotations of the eye that take the point of gaze from one region of interest to the next. Saccades are the principal way of bringing images of selected objects to the fovea, where visual acuity is best. In most everyday life tasks, people make saccades in sequence from one selected object to another, allowing gaze to pause briefly at different portions of a visual scene. These brief pauses between saccades are sometimes referred to as *fixations*. Figure 1 shows an example of a sequence of saccadic eye movements made while counting a set of dots. The upper portion of the figure shows the position of the line of sight superimposed on the display of dots. The lower portion shows the identical pattern of eye movements in a different format, namely, as horizontal and vertical eye positions as a function of time. The rapid shifts in eye position are the saccades. Notice that the pauses between the saccades last only about one half to one third of a second.

3. *Vergence eye movements.* The visual fields of the two eyes overlap, providing a basis for binocular and stereoscopic (three-dimensional [3-D]) vision. Vergence eye movements are movements of

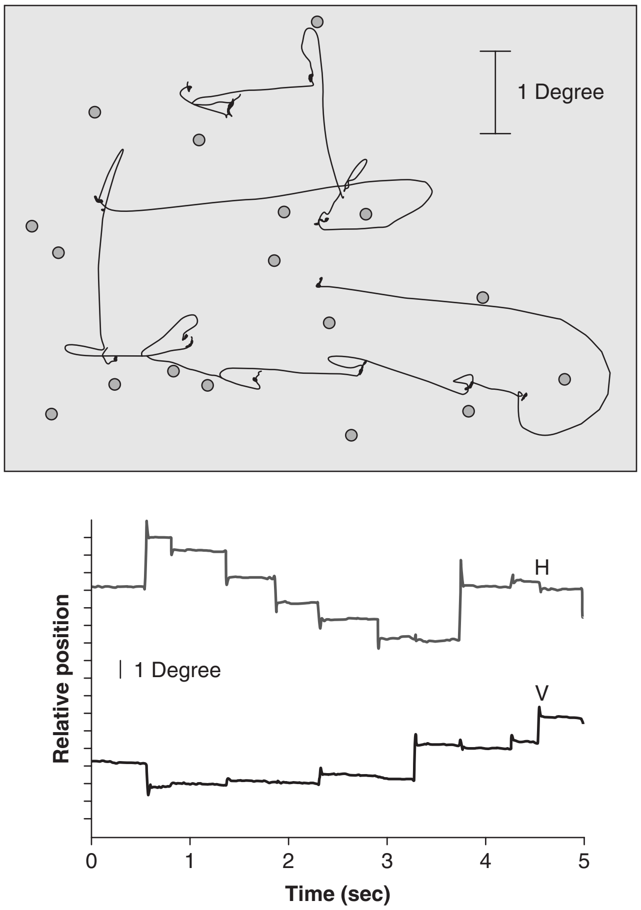


Figure 1 Sequence of Saccadic Eye Moments

Source: Wilder, J. D., Kowler, E., Schnitzer, B. S. Gersch, T. M., & Dosher, B. A. (in press). Attention during active visual tasks: Counting, pointing or simply looking. *Vision Research.*

Notes: (Top) Eye movements while counting dots. The line of sight began near the center of the display. The first saccade was made down and to the right. The next five saccades stepped across to the left, followed by four more saccades in the upper portion of the display. (Bottom) Horizontal and vertical eye positions over time. The rapid shifts in position are the saccades.

the eyes relative to each other that control the spatial registration (*correspondence)* of the two retinal images. Static vergence keeps the two eyes aligned, and dynamic vergence adjusts the angle between the lines of sight during binocular shifts of gaze between objects at different distances.

In addition to these three main types of eye movements, two other eye movement behaviors, maintained fixation and smooth pursuit, occur when we need to maintain a particular image on the fovea for extended periods. Both of these share characteristics with the types of eye movements listed, but each has special features that warrant a separate discussion. This entry discusses these three types of eye movements.

Compensatory Eye Movements

Rapid motion of the image on the retina produces blurring and smearing that will degrade vision. If the eyes were locked in their orbits, image motion on the retina would routinely reach unacceptably high levels with even the slightest movement of the head. The first line of defense against excessive retinal image motion is the vestibulo-ocular reflex (VOR). The vestibular organs (semicircular canals and otolith organs) sense motions of the head and relay this information to the oculomotor muscles. An approximately equal and opposite eye rotation is generated after a latency of only .008 seconds so that the direction of gaze is largely maintained. The VOR is supplemented by the slower optokinetic system (latency .08 seconds) that detects the global motion of the retinal image and elicits a smooth ocular following response. Compensatory eye movement systems, particularly the VOR, monitor their own effectiveness. They sense the amount of motion of the retinal image and alter the magnitude of the response by means of an adaptation process that maintains retinal image motion at a level optimal for seeing.

What is the optimal level of image motion for seeing? Excessive amounts of motion will degrade vision because of the image blur and smear. Conversely, a perfectly stationary retinal image is also undesirable: Experiments that have used optical methods to compensate for eye movements, so the image remains stationary on the retina, have shown that stabilized retinal images fade from view. Even slowly moving images are perceived poorly. To estimate the effectiveness of compensatory systems, and the resulting motion of the retinal image, Robert Steinman, Han Collewijn, and colleagues developed instrumentation to measure the movements of the eyes and head in three dimensions while people sit, stand, or engage in everyday life tasks, such as reaching or pointing. Their studies showed that the retinal image motions typically ranged from about 0.5 to 5 angular degrees/second. Interestingly, separate studies of effects of retinal image motion on vision showed that vision is usually best with image motions falling in this range.

Saccadic Eye Movements

Saccadic eye movements are rapid jumps of the eye used to bring the line of sight from one object of interest to the next. They can occur as often as three times per second and are found when we read, perform most any visuomotor activity, or casually inspect the environment.

Saccades can be made voluntarily, as can be demonstrated by deliberately looking from one side of a scene to another. Despite this voluntary character, most saccades occur without overt effort or awareness. For example, people are often surprised to discover that when reading the line of sight is continually jumping across the text, rather than moving smoothly. The lack of awareness of our own saccades is comparable with that of other motor activities. While walking, for example, a person can follow a specified manner of stepping or a marked trajectory, but such effortful performance is not natural. Normal walking is automatically guided by the layout of the surroundings and the level of urgency. Likewise, in natural situations, saccades are generated effortlessly, guided by the interest in exploring different portions of the environment.

Saccades are produced by brief, intense activity of the oculomotor muscles. Saccades typically last from .02 to .2 seconds, with saccade duration increasing with size. In natural activities, saccades are accompanied by rotations of the head. Early work on head-eye coordination led to the belief that during gaze shifts, the VOR automatically reduces the amount of eye rotation in the orbit by an amount equal to the head rotation, thus ensuring that gaze will land on the selected target regardless of how much the head moves. Recent work by Julie Epelboim and colleagues did not support this simple model and showed that the contribution of the VOR was constantly changing depending on whether the task required high levels of gaze accuracy or high gaze shift speeds.

Perceptual Consequences of Saccades

The frequent occurrence of saccadic eye movements is necessary for most visual tasks, but also presents formidable challenges to the visual system. As a result of saccades, the visual signal reaching the brain consists of a sequence of brief glimpses, lasting less than a half second, each centered on a different location (see Figure 1). Remarkably, subjective visual experience betrays none of this sequential quality. The world we perceive is stable and clear, with no perceptible jumps, interruptions, abrupt shifts of position, or awareness of the motion of retinal image during the saccades themselves. Three processes contribute to creating the impression of a clear and stable world.

1. *The suppression of vision during the saccade.* This phenomenon, termed *saccadic suppression,* refers to the inhibition of the visual signals generated during a saccade. Although many processes contribute to saccadic suppression, the major factor responsible is intra-saccadic visual masking. In intra-saccadic masking, the visual messages derived from the relatively stationary retinal images during the fixation pauses can suppress the messages generated from the rapidly moving image during the saccade.

2. *Nonvisual (extraretinal) signals representing saccadic commands.* Hermann von Helmholtz, working in the 19th century, proposed that during normal saccades a signal representing the saccadic command ("effort of the will," later termed *efference copy* or *corollary discharge*) is sent to the visual system to allow it to distinguish motion on the retina produced by saccades from genuine motion of the visual scene. Modern researchers have supported and developed Helmholtz's idea. Helmholtz's theory explains why illusory motion is perceived when attempting to move an eye that is paralyzed as a result of neurological disorders. An accurate and timely signal that represents the position of the eye can account for the ability to determine the true location of objects as we look around—an ability that is crucial for controlling reaching and locomotion in natural tasks.

3. *Selective attention.* If it were possible to construct a detailed representation of the world during each brief pause between saccades, the visual system would be confronted with a daunting task every time gaze shifts from one location to another. Every detail recorded before each saccade would need to be spatially aligned with the corresponding detail after each saccade to build a representation of a single and stable visual environment. Fortunately,

such painstaking construction is not needed because we perceive only a small portion of the visual environment during any fixation pause. Powerful, ever-present attentional filters limit immediate visual experience to what is most critical for the task at hand. Only attended information is preserved across saccades, and often in an abstracted (rather than point-by-point) representation. The same attentional filters that control access to perception also are involved in choosing the goal location of each saccade, a highly efficient way to share limited processing resources between perceptual and motor operations.

Vergence Eye Movements

Vergence eye movements align the lines of sight of the two eyes. Vergence is particularly important when we look at an approaching object, or when we look between two objects at different distances. Vergence facilitates the fusion of the views seen by left and right eyes into a single coherent percept, as well as the use of binocular vision to perceive a 3-D world (*stereopsis*). When you look from a distant to a near object (say, from a chair across the room to a paper on your desk) the lines of sight converge (rotate inward). When you look back to the distant object, the lines of sight diverge (rotate outward). The vergence eye movements made to look between near and far objects are rapid—much like the saccadic shifts of gaze made to look from left to right across a room.

Maintained Fixation and Smooth Pursuit

The discussion of eye movements so far has emphasized what the eyes do when shifting gaze from one object to another, or when compensating for movements of the head. Two other types of eye movements, maintained fixation and smooth pursuit, use the same neural systems that control saccades or compensatory movements, but are found in special situations that are important for vision, so their characteristics are worth noting.

Maintained fixation refers to the eye movements that occur when trying to look steadily at a single, stationary object for many seconds on end. Fixation was studied avidly in the 1950s and 1960s using an instrument, the contact lens–optical lever (a tight fitting large contact lens with a mirror attached), that was especially suitable for recording small eye movements at high resolution. Studies of maintained fixation were sparked by interest in theories that related visual acuity to the dynamic aspects of receptor responses.

In the laboratory, maintained fixation is studied when movements of the head are restricted by devices such as chinrests or bite bars. Under such conditions, fixation is remarkably precise, with the line of sight remaining most of the time in a region with a diameter of only eight photoreceptors. Nicole Putnam and colleagues, using new advanced adaptive optics methods of observing the retinal image directly, confirmed this level of precision, and showed that the preferred retinal locus of the fixation target was slightly displaced from the location of maximum cone density.

The eye is not perfectly stable during maintained fixation. Slow oscillations of the eye are interrupted periodically (once every 2 seconds to 2 or 3 times per second) by small saccades (less than 0.2 angular degrees). Fixational eye movements provide sufficient motion of the retinal image to maintain the visibility of the target. When head movements are permitted, the motions of the image on the retina are considerably faster, and more suitable for maintaining visibility of visual objects of a variety of sizes, eccentricities and contrasts.

Smooth pursuit eye movements occur when tracking an object that is moving smoothly across space. For objects moving either at a constant velocity or in a simple repetitive trajectory, smooth pursuit can attain velocities almost equal to that of the target. By contrast, smooth pursuit of targets that move in complex trajectories is poor because of the time required to detect and respond to changes in the stimulus path. The oculomotor system has an ingenious way to compensate for these processing delays, namely, by using available cues in the environment to predict the future trajectory of the target motion and generate smooth movements in advance of expected changes to the target path. Thus, it is possible to maintain acceptable levels of smooth pursuit provided that the motion of the target can be predicted with reasonable accuracy. In the case of truly random target motions—such as when following an insect or prey that can generate evasive patterns of motion—no predictive strategy will succeed in generating an accurate response.

Eileen Kowler and Han Collewijn

See also Eye Movements: Physiological; Eye Movements and Action in Everyday Life; Eye Movements During Fixation

Further Readings

Carpenter, R. H. S. (1992). *Movements of the eyes.* London: Pion Press.

Epelboim, J. (1998). Gaze and retinal-image stability in two kinds of sequential looking tasks. *Vision Research, 38,* 3773–3784.

Erkelens, C. J., Van der Steen, J., Steinman, R. M., & Collewijn, H. (1989). Ocular vergence under natural conditions. II. Gaze-shifts between real targets differing in distance and direction. *Proceedings of the Royal Society of London B, 236,* 441–465.

Gersch, T. M., Kowler, E., & Dosher, B. (2004). Dynamic allocation of attention during sequences of saccades. *Vision Research, 44,* 1469–1483.

Heinen, S. J., & Keller, E. L. (2003). Smooth pursuit eye movements: Recent advances. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (Vol. 2, pp. 1402–1414). Cambridge: MIT Press.

Kowler, E. (Ed.). (1990). *Eye movements and their role in visual and cognitive processes.* Amsterdam: Elsevier.

Putnam, N. M., Hofer, H. J., Doble, N., Chen, L., Carroll, J., & Williams, D. R. (2005). The locus of fixation and the foveal cone mosaic. *Journal of Vision, 5,* 632–639.

Steinman, R. M. (2003). Gaze control under natural conditions. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (Vol. 2, pp. 1339–1356). Cambridge: MIT Press.

EYE MOVEMENTS: EFFECTS OF NEUROLOGICAL AND MENTAL DISORDERS ON

Eye movement recordings provide powerful information about the decisions that a person's brain has made about whether something is important. Eye fixations cause the image of an object to fall on the fovea, the cone-rich high-acuity area of the retina, and so provide visual information at a high spatial resolution. The richness of the data provided by eye tracking is revealed in the plethora of dependent measures it provides. Moreover, the neural substrates of eye movement measures are largely known. Therefore, measurement of eye movements may provide insight into the underlying neuropathology of disorders affecting eye movements. Finally, because eye tracking is noninvasive, and tasks can be made simple yet still provide extremely rich data (e.g., participants can merely look at pictures of faces and natural scenes), the methodology is quite well suited to conducting critical developmental studies in young children. This entry describes how the visual scanpaths (the pattern of eye movements that occur when an individual processes a complex stimulus) of affected individuals are altered in some neurodevelopmental and psychiatric disorders. It highlights what the observation of altered scanpaths can reveal about underlying mechanisms in these disorders at the cognitive and neural levels of analysis.

Several basic types of eye movements either shift or stabilize gaze. Movements for shifting gaze direction include saccadic movements, which are rapid, ballistic movements of both eyes in the same direction; smooth pursuit movements that carefully follow a moving object; and vergence movements, which converge or diverge the eyes to focus appropriately on an object at various distances from the viewer. Gaze stabilizing movements include vestibulo-ocular movements (reflexes that stabilize images on the retina during head movements via the production of an eye movement in the direction opposite to the head movement) and the optokinetic reflex, which allows the eye to follow objects in motion when the head remains stationary. Most of the studies of eye movements in psychiatric and neurodevelopmental disorders have focused on movements that shift gaze. Disorders have included attention deficit hyperactivity disorder (ADHD), autism, schizophrenia, Tourette's syndrome, obsessive-compulsive disorder, anxiety, and depression. To illustrate this line of research, findings from studies of ADHD, schizophrenia, and autism are discussed here.

In an antisaccade task, an observer holds his or her fixation on a particular location, then a stimulus appears somewhere within the subject's visual field. The subject must move his or her eyes in the direction opposite of the visual onset. Success on this task requires inhibiting a reflexive saccade to the onset location, and purposefully moving the eye in the opposite direction. Neuroscience studies have revealed a critical role for prefrontal cortical regions in antisaccade tasks. An extensive body of research has demonstrated that ADHD is characterized by an increase of errors on antisaccade tasks (i.e., eye movements made in the wrong direction). These findings further characterize and quantify a cognitive mechanism underlying key aspects of the ADHD phenotype; those having ADHD have difficulties in response inhibition. Further, the work underscores abnormal development of the prefrontal cortex as an etiological factor.

Smooth pursuit movements have been studied extensively in individuals with schizophrenia. Such movements are under voluntary control in the sense that the observer can choose whether or not to track a moving stimulus. The brain systems involved in controlling smooth pursuit eye movements largely overlap with those involved in controlling saccades, and these movements work in an integrative way. The available studies consistently implicate reduced gain (ratio of eye velocity to the target velocity) as a characteristic marker in schizophrenia, meaning that their pursuit tends to lag behind moving (especially quickly moving) objects instead of smoothly and accurately "locking" onto and following the target as seen in neurotypical individuals. Of particular interest is that this attribute has been observed across children at risk for developing schizophrenia, children and adolescents with schizophrenia (i.e., childhood onset schizophrenia), and adults with schizophrenia.

Moving beyond the study of basic eye movements, some studies have examined visual scanpaths. Scanpaths are observed by monitoring point-of-regard over time. The eyes carry out a saccadic movement, fixate or rest momentarily, and then jump to a new region of interest.

Studies of how individuals with autism look at faces exemplify the power of recording visual scanpaths. When presented with a face, typically developing adults direct most of their attention to the core features of the face (i.e., eyes, nose, and mouth) and spend little time on other areas. In stark contrast, individuals with autism do not look at the eyes or the other core features of the faces. Instead, those with autism scan the faces in a seemingly random fashion. Such findings suggest a failure to understand the significance of the features for social understanding and could represent a mechanism leading to a failure to develop the normal specialization of some brain regions for processing social information.

Some of the most exciting recent work involves the study of young children at risk for developing disorders such as autism, ADHD, or schizophrenia by virtue of having a first-degree relative with the disorder. Ample evidence indicates the importance of genetic factors in the etiology of each. In searching for the causes of and potential treatments for these disorders, vital work is now underway to determine whether disrupted eye movements appear before the onset of the full phenotype. Eye tracking could be combined with computational modeling such that recorded visual scanpaths can be compared with the scanpaths predicted by models of visual attention that compute the "visual salience" of specific components of the scene via algorithms that use assumptions about how the brain works with regard to the operation of "bottom-up" (e.g., images are all analyzed the same way based on characteristics such as pixel brightness or darkness) and "top-down" (e.g., images are analyzed differently depending on the meaning of the image) perceptual and cognitive mechanisms. By thus comparing what saliency models predict for fixations of a particular image, and how subjects actually fixate that image, researchers could evaluate the relative contributions that bottom-up and top-down information processing mechanisms make to guiding eye movements. These contributions can then be compared in subjects with and without particular brain disorders.

Kevin A. Pelphrey

See also Attention: Cognitive Influences; Attention: Disorders; Eye Movements: Behavioral; Eye Movements and Action in Everyday Life; Eye Movements During Cognition and Conversation

Further Readings

Clementz, B. A., & Sweeney, J. A. (1990). Is eye movement dysfunction a biological marker for schizophrenia? A methodological review. *Psychological Bulletin, 108,* 77–92.

Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience, 2,* 1–10.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59,* 809–816.

Merin, N., Young, G. S., Ozonoff, S., & Rogers, S. J. (2007). Visual fixation patterns during reciprocal social interaction distinguish a subgroup of 6-month-old infants at-risk for autism from comparison infants. *Journal of Autism and Developmental Disorders, 37,* 108–121.

Munoz, D. P., Armstrong, I. T., Hampton, K. A., & Moore, K. D. (2003). Altered control of visual fixation and saccadic eye movements in attention deficit hyperactivity disorder. *Journal of Neurophysiology, 90,* 503–514.

EYE MOVEMENTS: PHYSIOLOGICAL

The *physiology of eye movements* refers to the brain regions and the connections between nerve cells in the various parts of the brain that are responsible for controlling how we aim and move our eyes. Because vision starts at the retina and the retina is attached to the back of the eye, visual perception depends on the proper operation of these brain circuits for eye movements.

If you have experienced how confusing it is to watch movies taken with a shaky camera, then you can readily appreciate how the ability to see and perceive the world depends on how the eyes are held and moved. Accordingly, one class of brain mechanisms, present in a wide variety of species, is responsible for keeping the eyes stabilized with respect to the visual world.

There are additional challenges for vertebrates such as primates, whose retina contains specialized regions (the fovea) that can be used to see a more limited portion of the visual scene at a much higher resolution. To make the best use of our visual system, we use a second class of mechanisms to move our eyes with respect to the visual world. These movements act to place and keep the images of interesting things within this specialized region of the retina, where they can be best seen. This entry describes physiological aspects of eye movements.

Behavior

Maintaining the Line of Sight

Two types of eye movements maintain the line of sight: the vestibulo-ocular reflex (VOR) and the optokinetic response (OKR). The VOR uses signals from the bony labyrinth of the inner ear that indicate how the head is moving in space. A direct three-neuron pathway relays these signals quickly (in less than 1 one-hundredth of a second!) to counter-rotate the eyeballs to keep the retinal images of the world stable during head movements. For example, as you look at this text, try shaking your head back and forth, and notice that your eyes automatically and rapidly move in the opposite direction, so that your ability to see the text remains unperturbed. As your head moves to the right, sensory neurons for the right inner ear increase their activity, whereas those in the left ear decrease their activity. These changes in activity are relayed by inter-neurons in the brain stem to the motor neurons that control the six muscles that move each eye. In addition, a copy of the signals from the inner ear are summed over time and stored in a process akin to mathematical integration; this integrated signal is also applied to the eye motor-neurons, so that your eyes are held in position when you stop moving your head.

The OKR also stabilizes the eyes, but uses visual signals about the movement of images across the retina. Because sensory transduction takes longer for vision than for the vestibular system, the OKR reacts slower than the VOR does (after 100 milliseconds rather than 10 milliseconds). However, unlike the signals from the inner ear, which decay after a few seconds of head movement, visual signals persist as long as images move across the retina. Thus, the OKR can compensate for low speeds and constant motions, even when there are no signals detected by the inner ear. For example, as you look at this text, try holding your head still and slowly move the page; the OKR will move your eyes to automatically follow the motion of the page, so that your ability to see the text again remains intact. The signals for this behavior come from areas of the brain that are specialized for processing visual motion. Some of these areas are located in the mantle of the cerebral cortex that covers most of the brain in primates, and others are located in nuclei underneath the cortex. Neurons in these areas provide inputs to the same inter-neurons and motor neurons in the brain stem that are used for the VOR. This shared brain stem circuit for holding and moving the eyes, which is used by all eye movement systems, is referred to as "the final common pathway."

Changing the Line of Sight

Vertebrates with foveal vision use two primary mechanisms for changing the line of sight: saccades and smooth pursuit. Saccades quickly move the image of a part of the visual scene from an eccentric location on the retina to the center of the retina where it can be seen best. For example, as you read this text, you use a sequence of saccades to flick your line of sight from word to word across the page. Saccades can be guided by any modality, and they occur frequently—a typical person makes several saccades each second. Saccades are fast and brief (lasting only tens of milliseconds), and they follow a stereotyped movement profile that is determined by a specialized circuit in the brain stem referred to as the *saccadic burst generator*. Within this circuit, several classes of neurons work together to produce the "burst" of activity that is needed to quickly rotate the eyes, as well as the "step" of activity that keeps the eyes in place after each saccadic eye movement ends. The properties of this circuit are finely tuned by pathways involving another brain structure, the cerebellum, which is the same brain structure necessary to execute other fine motor skills, such as playing the piano.

Smooth pursuit eye movements slowly rotate the eyes to match the motion of the visual target, such as a person walking across the field of view, and thus help avoid the blurry vision that would occur if the image of the target were allowed to slide across the retina. Unlike saccades, smooth pursuit is slower (speeds of less than 50 degrees per second) and is continuously updated by feedback about how well the target is being tracked. Notably, pursuit typically cannot be generated in the absence of a moving stimulus, although motion sensed with modalities other than vision can be used. Consequently, the motor circuits for pursuit do not construct the eye movement commands from scratch, but instead rely on descending signals from areas of the cerebral cortex that define the motion of the target to be tracked. These descending signals about target motion reach the final motor circuits through several routes. The most prominent pathway passes from the cortex through a set of relay nuclei at the base of the brain stem to parts of the cerebellum, which then change the activity of the same neurons involved in generating the VOR. In this way, pursuit eye movements commandeer the brain circuits also used for the VOR, but substitute a target-motion signal generated by the cortex for the head-motion signal provided by the inner ear.

These circuits in the brain stem for changing the line of sight are controlled by signals from higher centers located on the roof of the midbrain and several areas in the cerebral cortex of the forebrain. These connections with higher centers form the functional bridge between perception and the motor control of eye movements.

Brain Regions

Superior Colliculus

One of the most important sources of descending control signals is the superior colliculus (SC), located on the roof (tectum) of the midbrain. Neurons in the SC are organized to form a retinotopically organized map of the visual world—adjacent points in visual space are represented by adjacent SC neurons. The SC is a primary source of the signals needed by the saccadic burst generator to construct the motor commands for saccades. Because the final motor command is a precisely timed "burst" and "step" of activity, constructing these saccade commands involves a transformation from the spatial (i.e., retinotopic) code in the SC to the temporal code needed to properly control the activity of motor-neurons. This property of the brain mechanisms becomes evident when neural activity is artificially increased by passing small amounts of electrical current into the SC. Such stimulation evokes a saccadic eye movement with a particular direction and amplitude that depends on the location in the SC map. Stimulating for longer periods does not cause bigger movements, but instead evokes a sequence of several saccades, all of similar direction and amplitude, often referred to as a "staircase" because the plot of eye position over time resembles a set of stairs.

Activity in the SC is related both to the eye movements themselves and to the stimuli and events that elicit them: Some neurons show mostly movement activity, others show mostly sensory activity, and many show combinations of both. Many SC neurons also exhibit activity related to the remembered or perceived location of targets in the environment. Recent studies have shown that many SC neurons are involved in the preceding steps of identifying and selecting the target as well as in preparing to make eye movements. Other studies suggest that activity in the SC may be part of the neuronal circuits that control visual attention.

Cerebral Cortex

Several areas in the mantle of the cerebral cortex overlying the brain are important for eye movements, the most prominent of which is the frontal eye fields (FEF). As in the superior colliculus, stimulating the FEF with electrical currents can artificially induce saccadic eye movements with particular directions and amplitudes. Stimulation of an adjacent area of the cortex—the frontal pursuit area—causes smooth pursuit eye movements and is the only cortical area where stimulation can directly cause pursuit. Damage to the FEF impairs the generation of eye movements, especially when there are distracting stimuli, but these effects are temporary; in contrast, combined damage to both the SC and FEF causes a permanent deficit in the ability to voluntarily move the eyes. Neurons in the FEF show a range of properties: Some mostly respond to visual stimuli, some are active only during eye movements, and many exhibit activity for both sensory and motor events. The visual neurons also appear to play a role in identifying salient visual stimuli even in the absence of eye movements. Moreover, stimulation of the FEF with currents too low to cause eye movements can mimic the effects of directing attention to a particular location in space.

Several other areas in the cerebral cortex are important for generating eye movements. The lateral intraparietal area (LIP) plays a major role in the control of visual attention as well as eye movements and is often described as containing a "salience map" for visually guided behavior. The supplementary eye fields (SEF) play a less direct role in eye movements, but appear to be important for tasks that involve more arbitrary rules (for example, "move your eyes in the direction away from the visual stimulus"). The middle temporal area (MT) and the medial superior temporal area (MST) are important areas of the visual cortex for extracting the visual-motion signals necessary to drive smooth pursuit eye movements, as well as to support the perception of visual motion.

Basal Ganglia and Thalamus

In addition to the descending control signals provided by the SC and areas of the cerebral cortex, important regulatory circuits involve the basal ganglia and thalamus. The basal ganglia exert a strong inhibitory effect on the activity of neurons in the SC. By transiently stopping this inhibition (called disinhibition), the basal ganglia act like a gate to allow signals for eye movements to flow through the SC and exert their effects on the motor circuits. The basal ganglia appear to be especially important for eye movements involving reward and motivation, and this role likely involves the same neurotransmitter (dopamine) and neurons implicated in Parkinson's disease.

Ascending pathways through the thalamus back up to areas of the cerebral cortex involved in eye movement control. These regions of the thalamus collect signals from many of the structures described earlier (SC, cerebellum, basal ganglia, brain stem) and convey this information to the cortex. The function of this feedback is not fully understood, but it appears to update the visual and motor processes that take place in the cortex. In particular, by providing copies of the outgoing motor commands, the corollary discharge signals provided by the thalamus may allow us to correct for the sensory consequences of our own eye movements. For example, even though the visual world sweeps across our retina during eye movements, we correctly perceive the world as stable; this ability to maintain perceptual stability during eye movements may depend on these feedback pathways.

Richard J. Krauzlis

See also Action and Vision; Attention: Cognitive Influences; Corollary Discharge; Motion Perception; Neural Recording; Vision; Visual Acuity; Visual Processing: Subcortical Mechanisms for Gaze Control; Visual Search

Further Readings

Carpenter, R. H. S. (1988). *Movements of the eyes* (2nd ed.). London: Pion Press.

Krauzlis, R. J. (2005). The control of voluntary eye movements: New perspectives. *Neuroscientist, 11,* 124–137.

Krauzlis, R. J., & Stone, L. S. (1999). Tracking with the mind's eye. *Trends in Neurosciences, 22,* 544–550.

Leigh, R. J., & Zee, D. S. (2006). *The neurology of eye movements* (4th ed.). Philadelphia: F. A. Davis.

Schall, J. D., & Thompson, K. G. (1999). Neural selection and control of visually guided eye movements. *Annual review of neuroscience, 22,* 241–259.

Sommer, M. A. (2003). The role of the thalamus in motor control. *Current opinion in neurobiology, 13,* 663–670.

Sparks, D. L. (2002). The brainstem control of saccadic eye movements. *Nature reviews. Neuroscience, 3,* 952–964.

EYE MOVEMENTS AND ACTION IN EVERYDAY LIFE

We are largely unaware of the importance of eye movements in gathering the visual information that is used by the brain to perform even the simplest everyday actions. Take, for example, the eye movements involved in making a sandwich. If we watch what the eyes are doing we will see that they make a series of rapid jumps from one place to another, every few hundred milliseconds, much more rapidly than we are aware of, gathering visual information and guiding the movements of the hands. These rapid shifts of the eye from one place to another are called saccadic eye movements and are one of the primary eye movements involved in gathering information from different locations in a visual scene. These movements are necessary because we cannot see with equal clarity across the entire visual scene, but instead must move our eyes so that the object of interest falls on the high-resolution foveal region at the center of the retina. Objects off the line of sight are seen much less clearly, so if we want to grasp an object, such as a knife, we direct the center of gaze toward the regions where the fingers make contact. Almost without exception, whenever we pick up an object in everyday life, we direct our eyes to the object to guide the grasp. Similarly, we almost always direct our eyes at the things we are most interested in at the moment. Thus, eye movements are a critical component of everyday visually guided behavior, as this entry describes.

Eye Tracking in Natural Behavior

Despite this importance, the study of eye movements and their role in visual perception and action has historically been restricted to simple controlled experimental conditions. Part of the reason for this has been the general strategy in experimental psychology of investigating the simplest components of perception and movements with the goal of understanding complex behavior in terms of its elemental components. Perhaps the more important reason, however, has been the difficulty of measuring eye movements in an unconstrained observer. The high optical power of the eye means that a tiny eye rotation translates into a significant change in the direction of gaze in the world, such as a change in gaze from one word to the next during reading. To measure such tiny rotations, early eye position measurement devices involved awkward and often painful attachments to the eye that necessitated stabilizing the head. Subsequent systems, such as electromagnetic eye coils and the dual Purkinje image tracker were more comfortable, but still required the head to be held fixed. However, during the last 10 to 15 years, the advent and rapid refinement of eyetrackers mounted on the head have allowed considerable freedom of movement. Michael Land built one of the first of these, and a recent version by Jason Babcock and Jeff Pelz, shown in Figure 1(a) (see also the color insert, Figure 24a), is completely portable, using power from a battery carried in a small backpack. The new eyetrackers allow the study of eye movements over extended tasks in natural settings. As well as providing eye position in the head, or in space, these trackers also provide a video from the subject's viewpoint, with eye position superimposed. This is particularly useful in interpreting the eye movements because it shows the stage of the task corresponding to a particular gaze point.

Task Dependence of Gaze Patterns

Since the development of portable eyetrackers, eye movements in a variety of natural settings have

Figure 1 The RIT Wearable Eyetracker, Showing the Scene Camera and Eye Camera Mounted on a Lightweight Glasses Frame

Source: Hayhoe & Ballard (2005).

Notes: (a) The backpack carries a battery and video camera, which tracks eye position relative to the scene the person is observing. (b) Composite image of a scene from the head-mounted video camera while the subject makes a peanut butter and jelly sandwich. The images from the camera were integrated over different head positions. The fixations are indicated by the white circles. Circle diameters reflect the duration of the fixation. See also color insert.

been investigated. We might have expected that in the natural world, gaze would roam randomly around the scene, stopping at salient points that stand out from the surroundings, such as bright regions, or regions of saturated color, or edges between regions. (Note that the periods when the eye is relatively stationary, in between the high velocity saccadic gaze shifts, are called fixations.) A large body of theoretical work by Laurent Itti and colleagues has attempted to predict fixation locations by analyzing the content of the image for regions that differ from the surrounds in this way. If subjects are inspecting still images of scenes, these saliency calculations can predict some of the fixations, but in natural behavior such as driving, walking, sports, and making tea or sandwiches, fixations are instead guided by the visual information that the subject needs at the moment. The central result from investigating natural behavior is that gaze locations are tightly linked to the task, and observers rarely look at regions or objects that are irrelevant to the task at hand, regardless of their image saliency. Figure 1(b) (color insert, Figure 24b) shows all the fixations made while making a sandwich. When viewing a video replay from the subject's viewpoint, with eye position superimposed, we can see that fixations are tightly linked to the actions, such as grasping and moving objects, moving on to the next object when the needs of the current action have been met. To understand this better, imagine a small segment of behavior, when the subject has finished making the sandwich and is about to cut the bread. Gaze will be directed at the knife handle to guide the hand to pick it up. As the hand closes on the knife, the eye will move to the corner of the sandwich where the knife tip will be placed to begin cutting. The gaze then moves slowly along the bread with the knife following shortly afterward. After the sandwich has been cut, the eye will move to a location on the table where the knife is to be placed. Thus, the basic structure of the task allows one to link the visual operations fairly closely in time with the occurrence of eye and hand movements.

The sequence of fixations is tightly linked to the task, and many of the fixations appear to have the purpose of obtaining quite specific information. For example, subjects fixate the middle of the jar for grasping with the hand in a vertical posture, and the rim for putting on the lid, with the hand in a horizontal posture. This suggests that the visual information being extracted controls the pre-shaping in one case and the orientation of the lid in the other. Dana Ballard and colleagues have called this a "just-in-time" strategy, where observers acquire the specific information they need just at the point it is required in the task. Interestingly, although subjects are in general aware that they are looking at the objects involved in the task, they are not aware of the speed and specificity of the eye movements, which occur more or less automatically. Moreover, when grasping a jar, for example, the subject might register size and location, but not be aware of the label on the jar or even the color if the contents. Thus, “seeing” is not a unitary process, despite our subjective impression that we effortlessly perceive everything at once. Rather, it is an active process and is inextricably linked to the observer’s behavioral goals.

Learning Where to Look

Implicit in much of the research on natural tasks is the finding that eye movement patterns must be learned. For example, in tea making, observers must have learnt what objects in the scene are relevant because almost no fixations fall on irrelevant objects. In driving, Hiroyushi Shinoda and colleagues showed that approximately 45% of fixations fell in the neighborhood of intersections. As a consequence, subjects were more likely to notice stop signs located at intersections rather than in the middle of a block. Thus, subjects have learnt that traffic signs are more likely near intersections. Subjects must also learn the optimal location for the information they need. For example, when pouring tea, fixation is located at the tip of the teapot spout. Presumably, flow from the spout is best controlled by fixating this location. Another important way in which eye movements reveal the influence of learning is that they are often anticipatory; that is, saccades are often made to a location in a scene in advance of an expected event. For example, in Michael Land and Peter McLeod’s investigation of cricket, batsmen anticipated the bounce point of the ball, and more skilled batsmen arrived at the bounce point about 100 milliseconds (ms) earlier than less skilled players. Anticipation is not limited to skilled sports players. Even in less skilled activities, such as catching a ball, subjects reveal similar anticipatory eye movements. Thus, eye movement patterns reflect learning of the dynamic properties of the world.

Another important and unexpected feature of eye movements in ordinary behavior is the reproducibility of the eye movement patterns and the similarity between observers. For example, when cutting a sandwich, subjects always fixate the initial point of contact with the knife and move their gaze along the locus of the cut, just ahead of the knife. Given the role of eye movements in gathering information to accomplish tasks, it makes sense that fixation patterns between individuals should be similar because they reflect the physical and dynamic properties of the environment. This stability in fixation patterns makes research on natural behavior unexpectedly accessible to experimental investigation.

Neural Mechanisms of Gaze Control

Advances in understanding the role of eye movements in everyday behavior have been paralleled by advances in understanding the neural mechanisms underlying these movements. It has become clear that internal reward mechanisms are intimately linked to the neural machinery controlling eye movements. Wolfram Schultz and colleagues have shown that dopaminergic neurons in the basal ganglia (an important mid-brain structure involved in the generation of movements) signal the reward expected from an action. Sensitivity to reward is manifest throughout the saccadic eye movement circuitry. The neurons involved in saccadic targeting respond in a graded manner to both the amount of expected reward and the probability of a reward, in the period before execution of the response. Sensitivity to both these variables is critical for learning and, consequently, for linking fixation patterns to task demands. Thus, at both behavioral and neural levels, the target selection mechanisms appear to be intimately related to behavioral outcomes.

Applications

A particular advantage of investigating eye movements in everyday behavior is that the outcomes are much more directly related to both applied and clinical problems than to work in simpler, more controlled experiments. For example, if we want to know the consequences for everyday life of visual losses from disease or injury, it is necessary to understand how information in the visual world is used for particular tasks, such as navigation, obstacle avoidance, or preparing a meal. There is a long history of investigation of reading, but more recently, investigators such as Kathleen Turano and Krystel Huxlin have examined gaze patterns in individuals with visual impairments in a wider range of tasks. Such work can reveal potential mechanisms for compensating for visual losses and indicate potential rehabilitation strategies. This is true for visual losses caused by central damage such as that caused by stroke, as well as losses from peripheral disease. The clinical applications of eye movements in everyday behavior are an important area of research that has been greatly facilitated by the development of more "user friendly" eyetrackers, so we can expect to see exciting developments in this area.

Mary Hayhoe

See also Attention: Object-Based; Attention: Spatial; Eye Movements: Behavioral; Eye Movements During Cognition and Conversation; Visual Memory

Further Readings

Droll, J. A., Hayhoe, M. H., Triesch, J., & Sullivan, B. T. (2005). Task demands control acquisition and storage of visual information. *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1416–1438.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences, 9*(4), 188–193.

Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature reviews. Neuroscience, 2,* 194–203.

Land, M. (2004). Eye movements in daily life. In L. Chalupa & J. Werner (Eds.), *The visual neurosciences* (Vol. 2, pp. 1357–1368). Cambridge: MIT Press.

Land, M. F., & McLeod, P. (2000). From eye movements to actions: How batsmen hit the ball. *Nature Neuroscience, 3,* 1340–1345.

Martin, T., Riley, M., Kelly, K., Hayhoe, M., & Huxlin, K. (2007). Visually guided behavior of homonymous hemianopes in a naturalistic task. *Vision Research, 47,* 3434–3446.

Patla, A. E., & Vickers, J. N. (1997). Where and when do we look as we approach and step over an obstacle in the travel path? *Neuroreport, 8,* 3661–3665.

Pelz, J. B., & Canosa, R., (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research, 41,* 3587–3596.

Triesch, J., Ballard, D. H., Hayhoe, M. M., & Sullivan, B. T. (2003). What you see is what you need. *Journal of Vision, 3*(1), 86–94.

Turano, K. A., Geruschat, D. R., & Baker, F. H. (2003). Oculomotor strategies for the direction of gaze tested with a real-world activity. *Vision Research, 43,* 333–346.

EYE MOVEMENTS AND READING

When we read, we have the impression that our eyes glide smoothly across the page. However, this is an illusion. The eyes stay relatively fixed for about 200 to 300 milliseconds (ms) and then quickly jump (in about 20–30 ms) to a new location. The stable periods are called *fixations* and the eye movements are called *saccades*. Vision is suppressed during saccades (to prevent the perception of blur during eye movements) and useful information is obtained only during fixations. Our eyes typically move forward through the text, although *regressions* (saccades going backward through the text) account for about 10 to 15% of the movements. When readers come to the end of a line, a large leftward eye movement (called a *return sweep*) is made to near the beginning of the next line of text.

It is now easy to record eye movements accurately. The most precise eyetrackers have a spatial accuracy of less than one half a character, and they provide a millisecond-by-millisecond record in the time domain. Eyetrackers that are cheaper and easier to use are still spatially accurate to about a character and to about 5 ms temporally. Thus, with the types of tracker being used today, one can be quite sure which word a reader is fixating and can measure the duration of a fixation with an accuracy of less than 5 milliseconds (ms). This precision is important because what the eyes are fixating on is closely linked to the underlying cognitive processes of reading, as this entry describes.

The Perceptual Span in Reading

The most direct way that a tight linkage between the location of the eyes and the underlying processing in reading has been demonstrated is through the *moving window* paradigm first used in classic experiments by George McConkie and Keith Rayner. In this paradigm, what the reader can process from text on the video monitor depends on where the eyes are. For example, in a one-word window condition, the word that is being currently fixated is normal, but all the letters in the other words are replaced by Xs or random letters. However, when the eyes move and the reader fixates a new word, that word is now normal on the

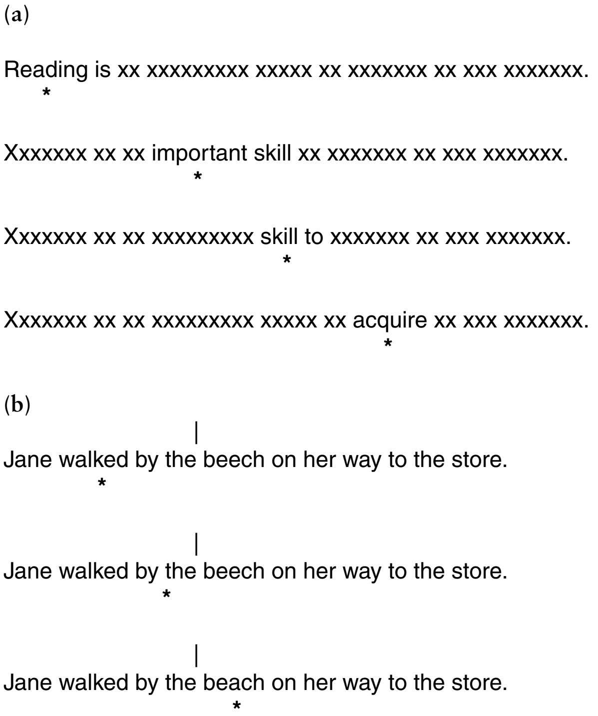


Figure 1 The Moving Window and Boundary Paradigms

Notes: (a) An illustration of the moving window paradigm. The size of the window in the illustration is two words: the fixated word indicated by the asterisks and the word to the right. As indicated in the text, the size of the window can be defined either by the number of letters or the number of words. In addition, the location of the window can be manipulated so that, for example, a two-word window could include the fixated word and the word to the left. (b) An illustration of the boundary paradigm. The vertical mark indicates the boundary (which is not visible to the participant). The display change involves only a single word location, and the only display change occurs when the boundary is first crossed. In this illustration, a homophone of the target word is the preview. Other possible previews could be bench, broad, bsfhfs, or xxxxx. Fixation time on the target word is shorter when "beech" is the preview rather than "bench" (indicating phonological coding of the preview), but "bench" is a better preview than any of the others, indicating orthographic coding of the preview.

next fixation and all the other words (including the previously fixated word) have their letters replaced by Xs (see Figure 1a). Thus, this paradigm is like a custom "slide show," where there is a new slide on each fixation (tailored to the fixation location). Provided that the window isn't less than a word, readers are generally not aware of the window; it just seems like normal reading.

The logic of the paradigm is that if all that was encoded on a fixation was the word being fixated, then readers would read normally with a one-word window. However, the reading rate in the one-word window condition is about 200 words per minute, which is about 60% of the normal reading rate. In a two-word window condition (in which the window contains the fixated word and the word to the right) reading is about 90% of normal, and in a three-word window condition (the window contains the fixated word and the two words to the right) reading is normal. Thus, the *perceptual span* (the region where readers extract information at least some of the time) appears to be three words (or more precisely, 3 to 4 letters to the left of fixation to about 14 to 15 letter spaces to the right of fixation). There is also nothing magic about the exact size of the window, however, because it is smaller if the text is quite difficult to decode. It's also smaller for beginning and dyslexic readers than for skilled readers, likely because they have more difficulty encoding the words in the center of vision than more skilled readers do. The asymmetry of the perceptual span observed in English is also not a hardwired "biological" constraint. In Hebrew, which is read from right-to-left, the perceptual span is asymmetric to the left. Thus, ongoing processing and visual attention modulate both the size and shape of the span.

The outer limits of the window to the right, however, are largely defined by visual acuity constraints: acuity is greatest in central vision (the fovea) and drops off quite quickly. However, it is best to think of the window as being defined by the number of characters instead of visual angle because there is a trade-off in acuity if the distance of the text or the size of the font changes. That is, when the text is larger, the letters are bigger (which makes them easier to process), but they are also further from the fovea, and those two factors generally compensate. Finally, the perceptual span for readers of languages that are more densely packed than English tends to be smaller than for English; for example, the perceptual span in Chinese extends about 1 character to the left of fixation to two to three characters to the right.

Eye Fixations Reflect Ongoing Cognitive Processing

The results of the moving window experiments described suggest that a major part of the agenda on a fixation is to encode the word that is fixated (though readers also get preview information from the next word). A host of evidence supports this point, most notably that characteristics of the fixated word—such as length, frequency in the language, and predictability from the prior context—affect the amount of time spent on the word. Word length, frequency, and predictability are highly correlated. Short words tend to be frequent and predictable, especially function words (articles, prepositions, conjunctions, and pronouns). Experiments, however, can unconfound these variables. For example, to examine word frequency, two words matched in length and predictability, but varying in frequency, would be inserted into the same sentence frame with the sentences closely matched on naturalness. A given reader would only read one of the two sentences, but in the counterbalanced design, equal numbers of readers would read each sentence.

These experiments indicate that the average fixation time on low-frequency words is 30 to 50 ms longer than those on high-frequency words. This difference does not merely reflect a few long times on low-frequency words; the difference is seen largely as a shift of the entire distribution of fixation times. Similar findings are observed for word length and predictability (when the other factors are equated); fixation times are shorter on predictable words than on unpredictable words and on shorter words than on longer words. These variables also influence the probability of skipping a word. Shorter, more frequent, and more predictable words are skipped more often; skipped words are processed on the fixation before the skip.

The discussion to this point has documented that the time to identify a word strongly influences how long the word is fixated and that these variables influence whether a word is skipped. This suggests that word identification is the primary engine driving the eyes forward in reading and is the process that largely determines reading speed. Rather direct evidence for this assertion comes from experiments in which the text disappears (or is masked) early in a fixation. If readers get to see the text for 50 to 60 ms before it disappears, reading proceeds quite normally because this is enough time for the visual information to be processed. Even though the word is no longer there, readers keep their eyes in place longer for low-frequency words than for high-frequency words!

The Processing of Words

It is sometimes suggested that skilled readers go directly from print to meaning without accessing the sound of the language. However, the involvement of phonological coding (i.e., translation to sound) is an important part of the process of identifying words. One demonstration of how early phonological coding enters into silent reading is with the *boundary paradigm,* which is similar to the moving window paradigm (but with only a single display change). An invisible boundary is set just to the left of a target word (see Figure 1b). Before the reader crosses the boundary, there is typically a preview different from the target word; when the eyes cross the boundary, the preview is replaced by the target word. The sentence makes sense with the target word, and readers are unaware of the identity of the preview and of the display change. Experiments using this technique have demonstrated that if there is letter overlap between the preview and the target, fixation times are shorter on the target word than if the preview shared no letters with the target. This indicates that partial information from the preview is extracted that facilitates processing of the word when it is fixated. However, this could either be orthographic (letter) information or phonological information. To test whether phonological information is involved in this process, experiments employed previews that were either homophones of the target word, for example, or control words that were matched on the number of letters they had in common with the target word (see Figure 1b). Fixation times on the target word were shorter when the homophone was the preview indicating that phonological coding occurs in early reading—even before a word is fixated.

Other properties of the fixated word besides length, frequency, and predictability also influence how long readers look at it. For example, the age when people first encounter a word yields a significant effect even when word frequency is controlled. In addition, whether a word has more than one distinct meaning (lexical ambiguity) and whether the prior context helps disambiguate the meaning of a lexically ambiguous word both influence how long readers look at the word. Besides ambiguity at the lexical and phonological level, ambiguities are also in written text at the level of syntax. Thus, a sentence such as, “While Mary was mending the sock fell off her lap,” is ambiguous because “the sock” can initially be taken as the object of “was mending” (an incorrect parse of the sentence) or as the beginning of a new clause (a correct parse of the sentence). Numerous studies using syntactically ambiguous sentences have examined readers’ parsing strategies, and the resulting eye movement data have been informative. It appears that readers usually construct only one of the possible syntactic structures, so that when the actual parse of the sentence is not what they expect, they often make either long fixations or regressions back to where they would have been led astray (or both). Thus, instead of computing all possibilities (which may overload memory), readers appear to be committed to one interpretation, but when they encounter information that contradicts that interpretation, they revise their initial parse.

The global meaning of a passage can also influence eye movements. Thus, readers often need to engage in anaphoric reference (e.g., tying an anaphor such as “he/she” to a prior antecedent). When there is ambiguity about what the antecedent is, fixation times increase and readers are also sometimes forced to make a regression back in the text. Likewise, clever experimental manipulations have revealed that when readers make inferences as they read, the nature of the inference can be inferred from the pattern of eye movements.

Models of Eye Movements During Skilled Reading

A number of computational models that account for the details of eye movements during reading have arisen to account for what has been learned about eye movements in reading. These models differ on some key dimensions, but they all do a reasonably good job of accounting for the findings described earlier. Although they have concentrated on the relationship between eye movements and word recognition rather than on higher level processes (such as syntax), the development of these models has been important in aiding our understanding of the relationship between eye movements and cognition during reading.

Keith Rayner and Alexander Pollatsek

See also Eye Movements: Behavioral; Eye Movements and Action in Everyday Life; Eye Movements During Cognition and Conversation

Further Readings

Andrews, S. (Ed.). (2006). *From inkmarks to ideas: Current issues in lexical processing.* Hove, Sussex, UK: Psychology Press.

Crowder, R. G., & Wagner, R. K. (1992). *The psychology of reading.* Oxford, UK: Oxford University Press.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye movement control in reading: Comparison to other models. *Brain and Behavioral Sciences, 26,* 445–476.

Snowling, M. J., & Hulme, C. (2005). *The science of reading: A handbook.* Oxford, UK: Blackwell.

EYE MOVEMENTS DURING COGNITION AND CONVERSATION

Femme fatales, interrogators, and cult leaders know the importance of eye movements. Gaze can act as a signal of attraction, a window into someone’s thoughts, and a means of persuasion. The close link between eye movements and psychological processes results from the anatomy of the eyeball. Detailed information can only be discerned from a small region in the center of the field of view, the fovea. So, unlike a wide-angled closed-circuit television camera that can slowly and passively sweep a large area, the small fovea must be rapidly and constantly moved around the world. For these motions to be efficient, they have to be intelligent. Therefore, eye movements are closely linked to the beliefs, expectations, and intentions of the observer. This link between eye and mind has recently been exploited by cognitive psychologists, but, perhaps implicitly, we are all attuned to the gaze patterns of those around us and what they indicate about the observer’s mental state. This entry examines the connection between gaze and cognition, and the role this plays in social interaction and conversation.

The mental processes of a simple decision can be seen in how the eye moves between choices. For example, the distribution of eye movements among a set of faces will reveal relatively how much each is preferred. If someone is asked to make a simple judgment, such as placing a sound or an object into one of two categories presented on screen, that person's eye movements between the options will reflect how certain he or she is of each item's membership. Moreover, influencing how the eyes move can influence mental processes. Researchers recorded the eye movements of participants interpreting an ambiguous picture or solving a difficult deductive problem from a diagram. Using low-level visual cues, such as brightening some areas of the image, a second set of participants were then influenced to attend to the same regions of the picture. The second set of participants was more likely to form the same interpretation of the ambiguous picture, and remarkably, was more likely to solve the deductive problem.

We do not have to be looking at objects or pictures for gaze to reveal cognitive processes. When looking at a blank screen and recalling an image or a painting, the eyes tend to move as they did during the initial viewing. When recalling nonvisual, spoken facts, the eyes will return to areas of a blank screen that were associated with the information. Not just memory, but imagery too can drive eye movements. Gaze becomes more active during spatial mental imagery and vivid dreams, and the eyes will move either horizontally or vertically when listening to a story about a train leaving the station or a man rappelling down a canyon.

How we move our eyes when listening to people speak reveals the time course of language comprehension and the way in which we ground spoken information in the visual world. For example, when people were asked to "pick up the candle" from the table, half way through the last word, they would often look at a candy lying nearby. This demonstrates that language comprehension is incremental: Listeners continually search for potential referents of spoken words in the visual world. Similarly, eye movements are yoked to language production. Speakers reliably look at objects just under a second before naming them. When two people talk to each while looking at the same visual scene, while watching a movie for example, their eye movements are closely coupled. The degree to which their eye movements are coordinated relates to how well they understand each other.

Eye movements during face-to-face conversation are governed by social forces and the process of coordinating dialogue. We can perceive each other's gaze accurately, and it is a powerful attentional cue. From infancy to adulthood, we use others' eye movements to infer their mental states. We are particularly sensitive to gaze that is directed at us. We can detect a millimeter displacement of a viewer's gaze away from the bridge of our own nose, which is at the limit of the visual acuity of the eye. Mutual gaze has two consequences. First, it correlates with many interpersonal factors, such as attractiveness, authority, and perceived sincerity. Second, the timing of mutual gaze is used to coordinate dialogue. Typically, a listener will look more at the speaker than vice versa. The speaker will periodically cast looks at the listener, creating a period of mutual gaze. At this point, the listener is likely to respond with a nod or "mhm," prompting the speaker to continue. In this way, eye movements are employed as part of the composite signal used in the collaborative act of conversation.

Too much can be made of gaze patterns, however. Neuro-linguistic programming holds that eye movements in certain directions correspond to the activation of certain brain regions or types of mental representation so that, for example, one can tell when someone is lying by the direction that he or she look. The wealth of scientific data discrediting neuro-linguistic programming is, however, matched by the volume of seminars and training courses promoting its use.

Uniquely poised between perception and cognition, eye movements are an invaluable tool for psychologists. Compared with the single data point provided by a button-press reaction time, the eye movements of a subject can provide researchers with a rich, dynamic data source concerning the temporal dynamics and psychological processes that led up to the response.

Daniel C. Richardson

See also Attention: Cognitive Influences; Decision Making, Perceptual; Eye Movements and Action in

Everyday Life; Language; Top-Down and Bottom-Up Processing

Further Readings

Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze.* Cambridge, UK: Cambridge University Press.

Henderson, J. M., & Ferreira, F. (Eds.). (2004). *The integration of language, vision, and action: Eye movements and the visual world.* New York: Psychology Press.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination: Common ground and the coupling of eye movements during dialogue. *Psychological Science, 18*(5), 407–413.

EYE MOVEMENTS DURING FIXATION

As you read this, your eyes are rapidly flicking from left to right in small hops, bringing each word sequentially into focus. When you stare at an object, your eyes will similarly dart here and there, resting momentarily at one place on the object, and then moving to another. But these large eye movements, called saccades (see color insert, Figure 25a), turn out to be just a small part of the daily workout your eye muscles are getting. Your eyes *never* stop moving, even when they are apparently settled, say, on a person's nose or a sailboat bobbing on the horizon. When the eyes fixate on something, as they do for 80% of your waking hours, they still jump and jiggle imperceptibly in ways that turn out to be essential for seeing. The tiny eye motions that you produce whenever you fixate your gaze are called fixational eye movements (see color insert, Figure 25b). If you could somehow halt these miniature motions while fixating your gaze, a static scene would simply fade from view. This entry discusses neural adaptation, visual fading, and microsaccades,

Neural Adaptation and Visual Fading

That the eyes move constantly has been known for centuries. In 1860, Hermann von Helmholtz pointed out that keeping one's eyes motionless was a difficult proposition and suggested that "wandering of the gaze" prevented the retina from becoming tired.

Animal nervous systems may have evolved to detect changes in the environment because spotting differences promotes survival. Motion in the visual field may indicate that a predator is approaching or that prey is escaping. Such changes prompt visual neurons to respond with neural impulses. Unchanging objects do not generally pose a threat, so animal brains—and visual systems,—did not evolve to notice them. Frogs are an extreme case because they produce no spontaneous eye movements in the absence of head movements. For a resting frog, such lack of eye movements results in the visual fading of all stationary objects. Jerome Lettvin and colleagues stated that a frog "will starve to death surrounded by food if it is not moving" (1968; p. 234). Thus, a fly sitting still on the wall will be invisible to a resting frog, but once the fly is aloft, the frog will immediately detect it and capture it with its tongue.

Frogs cannot see unmoving objects because an unchanging stimulus leads to *neural adaptation.* That is, under constant stimulation, visual neurons adjust their gain so they gradually stop responding. Neural adaptation saves energy but also limits sensory perception. Human neurons also adapt to sameness. However, the human visual system does much better than a frog's at detecting unmoving objects because human eyes create their own motion, even during visual fixation. Fixational eye movements shift the visual scene across the retina, prodding visual neurons into action and counteracting neural adaptation. Thus, eye movements prevent stationary objects from fading away.

In 1804, Ignaz Paul Vital Troxler reported that precisely fixating your gaze on an object of interest causes stationary images in the surrounding region gradually to fade away. Thus, even a small reduction in the rate and size of your eye movements greatly impairs your vision, even outside of the laboratory and for observers with healthy eyes and brains (see color insert, Figure 25c).

Eliminating *all* eye movements, however, can only be achieved in a laboratory. In the early 1950s, some research teams achieved this stilling effect by mounting a tiny slide projector onto a contact lens and affixing the lens to a person's eye with a suction device. In this set up, a person views the projected image through this lens, which moves with the eye. Using such a retinal stabilization technique, the image shifts every time the eye shifts. Thus, it remains still with respect to the eye, causing the visual neurons to adapt and the image to fade away. Nowadays, researchers create this same result by measuring eye movements with a camera pointed at the eye. The cameras transmit the eye-position data to a projection system that moves the image with the eye, thereby stabilizing the image on the retina.

Around the same time, researchers characterized three different types of fixational eye movements. *Microsaccades* are small, involuntary saccades that are produced when the subjects attempt to fixate their gaze on a visual target. They are the largest and fastest of the fixational eye movements, carrying an image across dozens to several hundreds of photoreceptors. *Drifts* are slow meandering motions that occur between the fast, linear microsaccades. *Tremor* is a tiny, fast oscillation superimposed on drifts. Tremor is the smallest type of fixational eye movement, its motion no bigger than the size of one photo-receptor.

Microsaccades in Visual Physiology, Perception, and Cognition

Since the late 1990s, fixational eye movement research has focused on microsaccades. Physiological experiments found that microsaccades increase the firing of neurons in the visual cortex and lateral geniculate nucleus by moving the images of stationary stimuli in and out of neuronal receptive fields. Perceptual experiments showed that the visual fading phenomenon described by Troxler is caused by the reduction of microsaccades that occurs during precise fixation.

Recent studies have shown that microsaccade rates are modulated by attentional shifts. Microsaccade directions may also be biased toward the spatial location of surreptitiously attended targets.

Fewer studies have addressed the neural and perceptual consequences of drifts and tremor. However, all fixational eye movements are likely to contribute significantly to visual perception. Their relative contributions may depend on stimulation conditions. For example, receptive fields near the fovea may be so small that drifts and tremor can maintain vision in the absence of microsaccades. Receptive fields in the periphery may be so large that only microsaccades are large and fast enough, compared with drifts and tremor, to prevent visual fading, especially with low-contrast stimuli. But it is also possible that if drifts and tremor could be eliminated altogether, microsaccades alone could suffice to sustain central vision during fixation.

Susana Martinez-Conde

See also Eye and Limb Tracking; Eye Movements: Physiological; Vision; Vision: Temporal Factors; Visual Illusions; Visual Processing: Primary Visual Cortex; Visual Processing: Retinal; Visual Processing: Subcortical Mechanisms for Gaze Control

Further Readings

Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., & Pitts, W. H. (1968). What the frog's eye tells the frog's brain. In W. C. Corning & M. Balaban (Eds.), *The mind: Biological approaches to its functions* (pp. 233–258). New York: Wiley.

Martinez-Conde, S. (2006). Fixational eye movements in normal and pathological vision. *Progress in Brain Research, 154,* 151–176.

Martinez-Conde, S., & Macknik, S. L. (2007). Windows on the mind. *Scientific American, 297,* 56–63.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience, 5,* 229–240.

Martinez-Conde, S., Macknik, S. L., Troncoso, X., & Dyar, T. A. (2006). Microsaccades counteract visual fading during fixation. *Neuron, 49,* 297–305.

Otero-Millan, J., Troncoso, X., Macknik, S. L., Serrano-Pedraza, I., & Martinez-Conde, S. (2008). Saccades and microsaccades during visual fixation, exploration and search: Foundations for a common saccadic generator. *Journal of Vision, 8*(14), 1–18, 21.

Troncoso, X., Macknik, S. L., Otero-Millan, J., & Martinez-Conde, S. (2008). Microsaccades drive illusory motion in "Enigma." *Proceedings of the National Academy of Sciences USA, 105,* 16033–16038.

Yarbus, A. L. (1967). *Eye movements and vision* (B. Haigh, Trans.). New York: Plenum Press.

EYES: EVOLUTION OF

Eyes detect the peak in the sun's electromagnetic radiation—light—and come in limited, convergent forms because of constraints set by the physics of light. Eyes are the sensory organs at the beginning of visual systems. A visual system such as our own, with its associated processing unit, is a complex of parts that have different functions and that are unrelated in their initial evolutionary history. A lens focuses light to form an image, retinal cells convert light waves to electrical signals, optic nerves carry those signals to the brain, the brain re-constructs the image and constructs color, and another part of the brain makes a decision about how the body will react to that image.

Eyes themselves can be classified as either chamber (or simple) or compound types. Chamber eyes have a single entrance for light, compound eyes have multiple entrances. But both produce an image on a retina—light sensors that do not cannot be branded eyes. The evolution of simple light sensors to become the first eyes during life's history holds hidden significance for eye evolution. This entry describes the evolution of the first eye, how the eye evolved, deficiencies of the eye, and the Cambrian explosion.

When Did the First Eye Evolve

For the first few hundred million years of multi-celled life on Earth, animals were blind, yet the selection pressures for vision were always immense. Holding back visual evolution was likely the high energy costs to build and maintain an eye.

That first eye, however, evolved in a trilobite (an arthropod predator, a distant relative of shrimps), around 524 million years ago (see Figure 1). Before that, animals possessed only simple light sensors. This is where the difference between an eye and a light sensor really matters. Light sensors provide information about night and day, or up and down in the water column. Eyes, on the other hand, allow animals to interact with each other using light as a stimulus. Eyes have evolved independently many times since, but the first eye to evolve introduced the sense of vision.

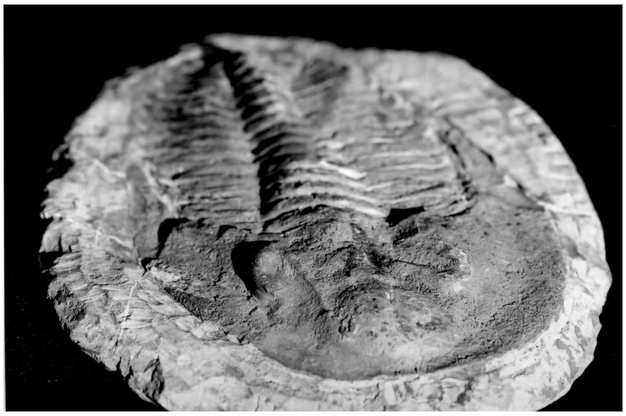


Figure 1 A *Cambropallas* Trilobite Fossil From Morocco

Notes: Around 524 million years old, frontal view. Eyes are protruding from the head shield (left eye is casting a shadow)—this was one of the first animals on Earth with eyes.

How the Eye Evolved

To understand the evolution of the eye, a computer model was developed that replicated the stages of evolution (Figure 2). This model showed

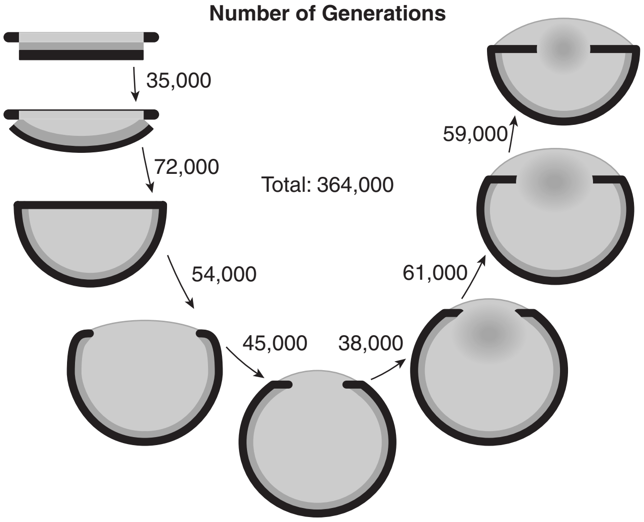


Figure 2 Dan-Eric Nilsson and Susan Pelger's Computer Model of the Evolution of a Chamber Eye

Source: From Land and Nilsson (2002).

Notes: A patch of light sensitive cells, with an outer (transparent) protective layer and a background dark, absorbing layer, evolves under a continuous selection for improved vision. The remaining parts of the visual system, however, do not necessarily evolve in parallel—imperfect vision can exist.

how a patch of light-sensitive skin could gradually transform into a chamber eye within 2,000 sequential changes, or just 364,000 generations (as a pessimistic estimate). The organ alters to first determine between sunlight-on or sunlight-off, then to sense the direction of sunlight, and finally to form an increasingly well-resolved image. The important point is that each stage confers an improvement in light perception or vision over its predecessor.

Recent research has centered on photoreceptor evolution based on molecular biological advances. Gene encoding opsin—the protein component of the light-sensitive visual pigment—and some developmental genes date back beyond the first eye and were recruited repeatedly during the evolution of different eyes, possibly between 40 and 65 times. Researchers can demonstrate the role of these developmental genes in the model organism *Drosophila*. By activating the developmental genes in different cells within a developing embryo, eyes will sprout on the legs and wings of the adult fly.

The evolution of lenses may owe something to physical phenomena such as molecular self-assembly or crystallization. If an enzyme, such as a liver enzyme, is overproduced within a cell, it will crystallize; our own lenses are formed this way. The phenomenon of crystallization enhances the effects of genetic mutation in the evolutionary process. Work underway aims to demonstrate this using the living cells that make the unusual calcitic lenses in brittle stars.

Deficiencies

There are signs of deficiencies—or what some might call imperfections—in all eyes. For example, in comparison with the retina of other animals, the human retina is inside out, with nerve fibers running back into the path of light rather than behind the retina (as they do in the eyes of squid).

Imperfections can also include failure to deliver on visual requirements. Such shortcomings can be revealed within the competition that exists between the eye of a predator and the color of its prey, for instance. Humans posses trichromatic vision (three types of cones in the retina, each responding to a different wavelength range), so we can see cuttlefish resting on the sea bottom, but cuttlefish predators cannot. Predators, in this case, are dolphins with monochromatic vision (they see in black and white because they have only one type of cone cell in their retinas), under which the cuttlefish are well camouflaged. A limitation of the human visual system is that we do not see ultraviolet light and, consequently, miss the ultraviolet patterns of flowers that attract insects.

Cambrian Explosion

Today, vision is the most universally powerful sense in its impact on animal interactions and behavior. With the evolution of the first eye, the size, shape, color, and behavior of animals were revealed for the first time—the position and movement of animals could be accurately tracked. Hence, the introduction of vision can be considered to be the launch of the most powerful weapon on Earth. And it evolved relatively quickly in evolutionary time. The evolutionary countermeasures sparked within the other, non-sighted, soft-bodied animals of the time were likely significant. Beginning immediately after this was the main event of the Cambrian explosion (approximately 540 million years ago), when animals simultaneously and independently evolved their hard parts. According to the light switch theory, the eye was the impetus for this event.

Since that first eye, vision has remained on Earth. Although only 6 of the approximately 37 animal phyla possess eyes, more than 95% of all species belong to these. Vision has been a powerful weapon and a successful innovation in the animal kingdom.

Andrew Parker

See also Animal Eyes; Evolutionary Approach; Vision; Visual System: Evolution of

Further Readings

Fernald, R. D. (2006). Casting a genetic light on the evolution of eyes. *Science, 313,* 1914–1918.

Land, M. F., & Nilsson, D.-E. (2002). *Animal eyes.* Oxford, UK: Oxford University Press.

Parker, A. R. (2003). *In the blink of an eye.* Cambridge, MA: Perseus Publishing.

Parker, A. R. (2005). *Seven deadly colours.* London: Simon & Schuster/Free Press.

EYEWITNESS TESTIMONY

Texas resident Larry Fuller was officially pardoned in 2007 for a rape he did not commit and for which he spent 19 years in prison—one of more than 200 erroneous convictions overturned by DNA evidence in the United States since 1989. In at least 75% of these cases, as in Fuller's, eyewitness error contributed to wrongful conviction. Partly to understand such errors, eyewitness scientists have built on 100 years of memory research in psychology to develop principles of perception and memory specific to the eyewitness experience. In the process, memory myths frequently held by both professionals and laypersons have been challenged. For example, research findings during the past three decades contradict the common belief that memories for traumatic events lie hidden in pristine and recoverable form. A more sophisticated and nuanced understanding has emerged—that any experienced event is perceived and encoded incompletely into memory and later recalled as an amalgam of true recollection and intruding non-memory elements. This entry describes issues relating to eyewitness testimony.

A police investigator might describe the perfect eyewitness as an observer who is attentive to all that transpires during a crime, who draws all meaningful crime details into memory, retains that information across time, and later recollects the crime with total accuracy. This ideal memory process—encoding, retention, and retrieval of experienced events—begins with perception. Before any information can be encoded into memory, an eyewitness's perceptual systems must be successfully deployed. Visual acuity, sharp hearing, and such properties of the senses, along with one basic requirement, namely that the witness is paying attention, will affect the quality of information available for encoding crucial details into memory. Unlike the ideal, however, actual eyewitnesses have less than perfect perception, a problem compounded by the complexities of the crime and the viewing environment (e.g., distance, illumination, visual obstructions, and perpetrator disguise). The limitations of human perception constrain the quality of resulting memory.

The eyewitness literature catalogues many perceptual difficulties for eyewitnesses, that is, factors that reduce the likelihood that useful, veridical details of a crime, including details about the perpetrator, are brought into and retained in the observer's memory. First, the eyewitness's attention to the event is crucial. The *weapon focus effect* exemplifies the intersection of crime context and perceptual processes. During a crime of short duration that involves a weapon, the eyewitness's attention is typically riveted on the threatening weapon that is suddenly central to his or her well-being. Later attempts to recollect peripheral details and recognize the perpetrator are less successful than for those eyewitnesses exposed to the same stimulus event without a weapon. This narrowing of visual attention can even interfere with the perception of auditory information. Also, witnesses tend to focus on action rather than persons, resulting in incomplete memory for the offender. Field demonstrations of *change blindness* further indicate that observers who concentrate on a central action or task fail to notice even large adjustments in the broader visual field. For example, 50% of persons who were asked by a stranger for campus directions and then unknowingly subjected to an experimental sleight-of-hand in which the stranger was surreptitiously replaced by a different person (as construction workers carried a large door between the subject and requester, temporarily blocking the subject's view of the stranger) failed to recognize the change even as they continued to provide the requested directions.

The short duration of many crimes also is disruptive to perception. Faces and events seen for only a brief time are difficult to scan completely or to perceive at a detailed level. Consequently, an offender of average appearance (absent a salient or distinguishing feature) may generate a nonspecific and blurry memory in the eyewitness. A perpetrator's disguise by hat, mask, or hood will reduce the facial-feature cues necessary for later identification, particularly if the hair or hairline is masked, as will a change in perpetrator appearance between event and later attempts at identification. Also, eyewitnesses are prone to distorted perception: Perpetrators are likely to seem taller and the duration of a brief crime longer than objective reality.

The stress of a crime will often constrain perceptual and encoding processes. Contrary to lay belief that a high level of stress will enhance memory ("I was terrified . . . his face is burned into my memory"), research in both lab and field demonstrates that although stress heightens alertness, high stress may disrupt effective encoding of details. A recent field test of military personnel in rigorous survival school interrogations documented dramatic reductions in eyewitness identification accuracy (later recognition of their interrogators) for high, versus low, stress conditions.

Individual witness characteristics may determine what the eyewitness pays attention to and the detail that is encoded. For example, a car buff may quickly discern the model of the get-away vehicle, or a gun expert may recognize the type of weapon. The high stimulus load of a complex crime event, criminal or civil, may be difficult to meaningfully encode by an eyewitness who is startled by or cognitively unprepared for such an occurrence. Similarly, an intriguing aspect of crime perception—and a principle that presumably rests on the past experience of the witness as well as the viewing opportunity—is the ability of eyewitnesses to more accurately perceive and later recognize the face of a same-race offender (the *own-race bias*). Finally, individual expectations about events influence how crimes are perceived. In general, we have a tendency to see what experience and beliefs have taught us to expect.

Scientists now understand that the content of perception—what the witness believes he or she saw at the time of the crime event—can be disrupted and changed through subsequent experience. Research has established beyond question that information from outside the true event can unwittingly fill memory gaps and rewrite memory, even for the face of the offender, a phenomenon dubbed the *misinformation effect.* More dramatically, whole episodic memories can be constructed from imagination in the absence of true experience. Indeed, laboratory participants have been led to believe that in childhood they were lost in a shopping mall for an extended time, broke a window and cut themselves, were attacked by an animal, or endured other upsetting experiences—"memories" in the absence of direct perception of a real personal event.

The trustworthiness of eyewitness testimony is further damaged by the fact that eyewitness confidence is distorted by factors external to true memory, for example, an investigator's confirmation that the eyewitness has chosen the police suspect from a lineup, or a therapist's affirmation of belief in the veracity of a client's recovered memory. Of particular relevance to perception processes is *the post-identification feedback effect.* An eyewitness who has received confirmatory feedback after a police lineup decision ("Good, you identified the suspect"), even if the choice is wrong, will report significantly more certainty about the identification than will a witness who did not receive feedback. Even more unsettling, the witness whose identification choice is confirmed will convey distorted retrospective memory for how the event *was perceived,* claiming a better view and greater attention paid to the perpetrator. The testimony of this eyewitness is likely to be quite believable at trial because he or she truly accepts this version of reality.

To improve the probative value of eyewitness evidence, scientists continue to investigate procedural strategies for interviewing and evidence-gathering, to protect the memory of the eyewitness and to keep the eyewitness's retrieved memory in close synchrony with original perception. Science-based eyewitness procedures help minimize future risk for the kind of tragedy that befell Larry Fuller and other victims of false identification and wrongful conviction.

Nancy K. Steblay and Elizabeth F. Loftus

See also Attention: Cognitive Influences; Attention and Memory; Face Perception; Vision: Cognitive Influences; Visual Memory

Further Readings

Doyle, J. (2005). *True witness: Cops, courts, science, and the battle against misidentification.* New York: Palgrave Macmillan.

Loftus, E. F. (2005). A 30-year investigation of the malleability of memory. *Learning and Memory, 12,* 361–366.

Wells, G. L., Memon, A., & Penrod, S. D. (2006). Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest,* 7(2), 45–75.

F

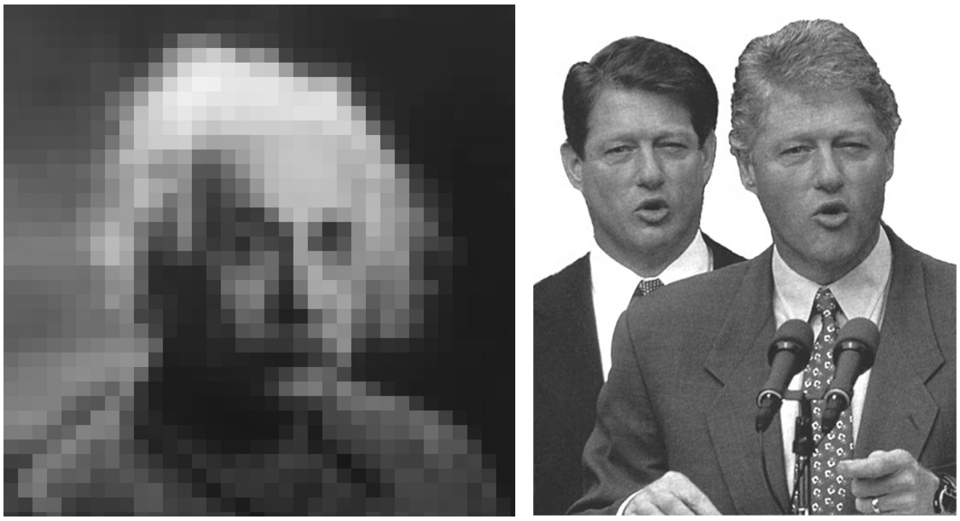
FACE PERCEPTION

Faces constitute one of the most pervasive and important classes of objects in our visual environment. The human visual system is remarkably proficient at various face perception tasks, including identification, emotion assessment, gaze estimation, and gender, race, and age classification. Much research effort has been directed toward elucidating the facial cues that underlie these different assessments, with identification attracting the lion's share. This entry provides an overview of four key aspects of human face perception performance: (1) tolerance for degradations, (2) relative contribution of geometric and photometric cues, (3) development of face perception skills, and (4) neural correlates of face perception.

Limits of Human Face Recognition Skills

The high resolution of foveal vision allows the human visual system, in principle, to discriminate between individuals on the basis of fine differences in their facial features. Analogously, progressive improvements in camera resolution make it possible to use increasing amounts of detail in face representations in machine vision systems. The advent of iris-based biometric systems, which use the fine details of iris patterns as identifying signatures, is a case in point. However, the problem that such details-based schemes often have to contend with is that high-resolution images are not always available. This is particularly true in situations where individuals have to be recognized at a distance. When seen from a distance, the image of a face occupies a small area of the retina, and thus is sampled by only a few photoreceptors. In effect, only the coarse low-resolution structure of distant faces is available for further processing. Precisely how does face identification performance change as a function of image resolution? This question is relevant for characterizing human face recognition performance, and for understanding the fundamental face representation strategies used by the visual system.

Working with block averaged images of familiar faces of the kind shown in Figure 1(a), researchers have found that observers can recognize faces well above chance levels with image resolutions of merely 7 × 10 pixels, and reach ceiling levels by resolutions of 25 × 30 pixels. The remarkable tolerance of the human visual system to resolution reduction demonstrates that fine featural details are not necessary to obtain good face recognition performance. Furthermore, given the indistinctness of the individual features at low resolutions, it appears likely that diagnosticity resides in their overall configuration. However, precisely which aspects of this configuration are important, and how we can computationally encode them, are largely open questions. Tentative clues to the answer come from experiments probing the relative saliencies of different facial features. Figure 1(b), for instance, demonstrates the perceptual importance of overall head configuration for face recognition. The internal features on their own and even their mutual configuration appear to be



(a) (b)

Figure 1 Stimuli for Face Perception Research

Sources: (a) Harmon, L. D., & Julesz, B. (1973). Masking in visual recognition: Effects of 2-dimensional filtered noise. Science, *180*(4091). (b) Sinha, P., & Poggio, T. (1996). I think I know that face . . . *Nature, 384*(404).

Notes: (a) Images such as the one shown here have been used by several researchers to assess the limits of human face identification processes. (b) Although this image appears to be a fairly run-of-the-mill picture of Bill Clinton and Al Gore, a closer inspection reveals that both men have been digitally given identical inner face features and their mutual configuration. Only the external features are different. It appears, therefore, that the human visual system makes strong use of the overall head shape to determine facial identity.

insufficient to account for the impressive recognition performance of subjects with full face images.

Cues for Face Recognition

For a cue to be useful for recognition, it must differ across faces and be consistent for a given face. Such a cue would be an *intrinsic* attribute of the face, such as its two- and three-dimensional shape, or pigmentation. A cue that differs indiscriminately, irrespective of facial identity, would be expected to be ignored by the human visual system. This would be an *extrinsic* cue, such as illumination. How perceptually important are these two sources of facial image variability?

Illumination greatly affects the image level appearance of a face. Surprisingly, humans appear to be unable to completely ignore illumination induced image changes while performing recognition tasks. When observers are asked to match different images of the same face, performance is worse when the two images of a face to be matched are illuminated differently. However, this confound is less pronounced for familiar faces, and naming performance for such faces is largely invariant across different illumination conditions, except for some unusual circumstances, such as lighting from below. Taken together, these results suggest that illumination-induced image changes are included in facial representations, but with increasing familiarity with a face (and, presumably, viewing it across many different lighting conditions), the contribution of illumination to recognition performance becomes less pronounced.

Besides lighting, two additional sources of variability for faces are their shape and their pigmentation. Given that we can recognize individuals even in unpainted stone busts, it is clear that shape cues on their own can support recognition. These anecdotal observations are supported by results from systematic studies with three-dimensional (3-D) laser-scanned faces that similarly lack variation in pigmentation. However, the ability to recognize a face in the absence of pigmentation cues does not mean that pigmentation cues are not used when available. Indeed, pigmentation may also be important for face recognition. Unlike other objects, faces are much more difficult to recognize from a line drawing than from a photograph, which suggests that the missing pigmentation cues may well be important.

Recently, some researchers have attempted to directly compare the relative importance of shape and pigmentation cues for face recognition. The basic approach is to create a set of face representations with the shape of a particular face and the average pigmentation of a large group of faces, and a separate set of face representations with the pigmentation of an actual face and the average shape of many faces. The rationale here is that to distinguish among the faces from the set that all have the same pigmentation, subjects must use shape cues, and vice versa for the set of faces with the same shape. Subjects are trained with one or the other set of face images and are subsequently tested for memory. Recognition performance is found to be about equal with both the shape and pigmentation sets, suggesting that both cues are important for recognition.

Evidence indicates that all sources of facial image variability affect the internal representations. When considering extrinsic cues such as illumination, this is somewhat of a drawback because a factor that is unrelated to facial identity ends up influencing performance. However, for the intrinsic cues of shape and pigmentation, the visual system behaves in an appropriately opportunistic fashion, making use of all cues that vary across exemplars of this class to achieve its impressive face recognition skills.

Development of Human Face Recognition Skills

Basic face perception skills develop surprisingly rapidly during an infant's development. As early as a few days after birth, infants appear to be able to distinguish their mother's face from other faces. Yet, after this initial phase of rapid progression in face recognition abilities, the full maturing of children's face-processing abilities is a long drawn out process, estimated to take more than a decade. This section outlines the trajectory of development of face processing.

Evidence From Newborns

As is the case for most visual skills, face processing must be connected with some primitive mechanism from which more advanced processes can be learned. A key question is are infants born with some innate abilities to process faces or are those abilities a consequence of general visual learning mechanisms? To examine this issue, newborns are assessed for various abilities as soon as is practical after birth. Three major findings have historically been taken as evidence for innate facial processing abilities: (1) the initial preference for faces over non-faces, (2) the ability to distinguish the mother from strangers, and (3) imitation of facial gestures.

Innate Facial Preference

Are infants pre-wired with a face detection algorithm? If infants knew how to locate the faces in an image, it would be a valuable first step in the subsequent learning of face recognition processes. John Morton and Mark Johnson formalized this idea into a theory called *CONSPEC and CONLERN.* CONSPEC is the structural information that guides newborn preferences for face-like patterns. CONLERN is a learning process that extracts further visual characteristics from patterns localized based on CONSPEC.

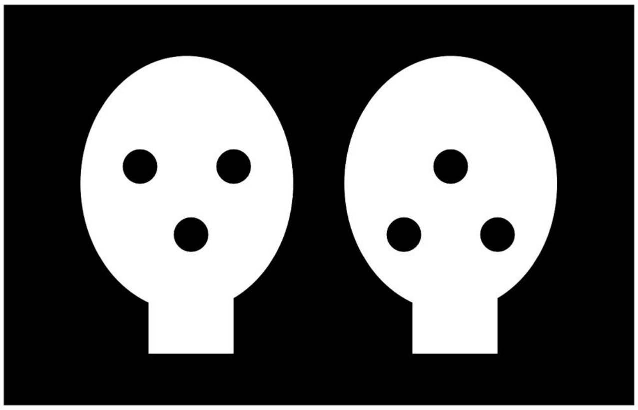


Figure 2 Newborn Gaze Orientation

Source: Johnson & Morton, 1991.

Notes: Newborns preferentially orient their gaze to the face-like pattern on the left, rather than the one shown on the right, suggesting some innately specified representation for faces.

Some supporting evidence for this theory comes from the fact that newborns do indeed preferentially orient their gaze to face-like patterns. When presented with a rudimentary drawing of a face, upright or inverted (Figure 2), newborns are reported to gaze longer at the upright pattern. The two patterns are largely identical except for their "faceness" and the infants have had little visual experience before the presentation, so this finding is interpreted as a demonstration of newborns' innate preference for faces.

Distinguishing the Mother From Strangers

Further evidence for innate mechanisms for facial processing comes from the remarkable ability of newborn infants within less than 50 hours of birth to discriminate their mothers from strangers. Newborns suck more vigorously when viewing their mother's face on a videotaped image. They also are capable of habituating to a mother's image, eventually preferring a novel image of a stranger, showing a classic novelty preference. External facial features, especially hair, are believed to underlie much of this discriminative performance of newborns.

Imitation of Facial Expressions

It has been reported that newborns are capable of imitating their caregivers' facial expressions. This would entail the infant's recognizing the expression, then mapping it to its own motor functions. Many studies have shown evidence of imitation. However, there have been some failures to replicate this result. Moreover, consistent positive results have been found primarily for one expression, namely, tongue protrusion. This action might be an innate releasing mechanism, perhaps to a surprising stimulus, or an attempt at oral exploration. Thus, the action might be a response to the stimulus, rather than a recognition of it followed by imitation.

Developmental Progression Beyond Infancy

Although infant face recognition proficiency develops rapidly within the first few months of life, performance continues to improve to the age of 10 years or even later. The best-studied progression is the use of featural versus configural facial cues.

Adults match upright faces more easily than inverted faces (the so-called inversion effect). This is believed to be a consequence of the disruption in configural processing in inverted faces. Interestingly, 4-month-old infants do not show this decrement in performance if the images to be matched are otherwise identical. However, the inversion effect does appear in 4-month olds if, for instance, pose is manipulated at each presentation of the face. Thus, evidence indicates that configural cues start to play some role in face processing at this early age, although the processing of these cues is still rudimentary. Processing based on features appears to play the primary role in infant facial recognition. Given the early onset of the usage of configural cues in child face recognition, rudimentary though it may be, one would expect that full maturation of such a fundamental system would ensue rapidly. However, numerous studies have found that, although face recognition based on features reaches near-adult levels by the age of 6 years, configural processing still lags behind until at least 10 years of age, with a gradual progression of greater accuracy and dependence on configural cues.

Neural Underpinnings of Human Face Perception

Whether faces constitute a "special" class of visual stimuli has been debated for many years. Since the first demonstrations of the inversion effect described previously, it has been suspected that unique cognitive and neural mechanisms may exist for face processing in the human visual system.

A great deal of evidence shows that the primary locus for human face processing may be found on the fusiform gyrus of the extrastriate visual cortex. This region shows an intriguing pattern of selectivity (schematic faces do not give rise to much activity) and generality (animal faces do elicit a good response), suggesting a strong domain-specific response for faces. In keeping with behavioral results, the fusiform face area (FFA) also appears to exhibit an inversion effect. Overall, the characterization of the FFA as a dedicated face-processing module appears strong. Developmental studies indicate a progressive strengthening of FFA activity with age. In a paradigm comparing the passive viewing of faces versus houses, younger children (8–10 years) showed weaker activation in the fusiform gyrus, relative to older (12–14 years) children. Thus, although the neuroimaging data are preliminary at the present, they do appear to be broadly consistent with the behavioral data showing continuing development of face processing well into adolescence.

The debate about faces being "special" is far from over. It has been suggested that rather than being a true "face module," the FFA may be responsible for performing either subordinate or expert-level categorization of generic objects. For instance, in dog experts, this cortical patch might be involved in discriminating one dog exemplar from another, whereas in a car expert, it might help distinguish different cars. The reason this area exhibits responsiveness to faces might be that most observers have attained expertise in face discrimination. There are results from both behavioral studies and neuroimaging studies that lend some support to this perceptual expertise account.

Neurophysiological evidence indicates that at least some face recognition tasks may be carried out over a surprisingly short time. Neurons in the primate inferotemporal (IT) cortex can exhibit selectivity to stimuli that are more complicated than the simple gratings and bars that elicit responses from cells in early visual areas. In particular, some cells in the IT cortex are selective for faces. Moreover, the latency of response in these cells is in the neighborhood of 80 to 160 milliseconds (ms). The computational relevance of these results is that recognition as it is performed to the level of the IT cortex probably requires only one feedforward pass through the visual system. Feedback and iterative processing are likely not major factors in the responses recorded in these studies.

Pawan Sinha, Yuri Ostrovsky,
and Richard Russell

See also Object Perception; Face Perception: Physiological; Perceptual Development: Face Perception; Visual Processing: Extrastriate Cortex

Further Readings

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science, 195*(4275), 312–314.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience, 4*(8), 2051–2062.

Harmon, L. D. (1973). The recognition of faces. *Scientific American, 229*(5), 70–83.

Harmon, L. D., & Julesz, B. (1973). Masking in visual recognition: Effects of 2-dimensional filtered noise. *Science, 180*(4091), 1194–1197.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302–4311.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198*(4312), 75–78.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN—A 2-process theory of infant face recognition. *Psychological Review, 98*(2), 164–181.

Sinha, P., & Poggio, T. (1996). I think I know that face . . . *Nature, 384,* 404.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141–145.

FACE PERCEPTION: PHYSIOLOGICAL

The investigation of the neural substrates of face perception constitutes an important area of study in both animal and cognitive neuroscience. Faces are used as stimuli in neurophysiological and neuroimaging studies for two primary reasons. First, from a practical standpoint, faces tend to evoke robust and strong responses relative to other objects, so that they are used in many studies that are not necessarily focused in the processing of faces per se, but their use may make the experiment more tractable. An example would be a study on the question of the fate in the visual system of information that is unattended by an observer: having the unattended stimuli be faces would allow experimenters to query known face-selective areas in the visual system for small responses under conditions of inattention. Second, an entire field of research is dedicated to the study of the perceptual and neural mechanisms that are specialized for face perception. This is the focus of this entry.

Only a few object categories have been argued to elicit category-specific responses in the human brain. These categories include places, words, letters, and body parts. Faces are of particular interest, both because of their critical importance for human social interactions and because face-specific responses can be studied across a number of nonhuman animal species, in particular in the monkey brain. In addition, because it is especially plausible that face perception may have evolved as a specialized dedicated system in the brain, the study of face perception plays an important role in the debate on the putative modularity of the mind.

A Specialized Mechanism for Faces

Whether faces are special or not is often mentioned as a topic of debate. Different aspects of this debate relevant to the physiological substrates of face perception are addressed later in this entry. But what elicits relatively little controversy is the body of evidence that shows a specialized response to faces in the visual system and elsewhere in the brain. In this context, "specialization" takes a descriptive meaning, highlighting the fact that the response to faces (in a neuron, in an entire brain area, or in a scalp recording) can be separated from the response to objects. Either the response is stronger for faces than for any other category tested (although typically faces are not the only stimulus eliciting a response), or the responses are spatially or temporally separated.

Face Cells in Animals

The evidence for a specialized response to faces comes from virtually every technique available to cognitive neuroscience. Single-cell recording studies in monkeys by Charles Gross and colleagues in the early 1970s revealed the existence of a small proportion of temporal lobe neurons that respond selectively to objects such as hands and face cells. Subsequent monkey studies revealed two populations of face cells, one in the inferotemporal cortex, which appears to code for the identity of faces, and one in the superior temporal sulcus, which is more sensitive to eye gaze and facial expression. Interestingly, face neurons have also been found in sheep, a social species that demonstrates recognition behaviors that are not unlike those of humans, such as an inversion effect (inversion affects the recognition of faces more negatively than it affects the recognition of other objects), when tested with both sheep and human faces. Whereas typical studies find only a small fraction of the neurons recorded from to be selective for faces, recent functional magnetic resonance imaging (fMRI) work in the monkey reveals that when patches that respond more to faces than to objects are identified to guide recordings, almost every neuron in these patches is face selective.

One question concerns how these populations of neurons may encode the information needed to support decisions about faces. One solution is that of a sparse code, with few neurons encoding the identity of each face for instance. At the extreme, this is referred to as the "grandmother" cell hypothesis, whereby only one neuron would be sufficient to code the identity of a given face. One advantage of such a coding scheme is a limited overlap between the representations of different faces, limiting interference on known faces when new faces are learned. A different solution is that of a distributed code, with the pattern of activity over a large population of cells representing the identity of a face. Relative to a sparse code, this scheme has more capacity (the number of faces that can be stored grows exponentially, rather than linearly, with the number of neurons in the population) and the representation of any one face is also more subject to damage. The current consensus is that a relatively sparse code is used, with a small number of cells that represent behaviorally relevant features (e.g., diagnostic of identity, expression, gaze direction). Recent work using linear classifier algorithms reveals that categorization and even identification of faces and other objects is robust based on the responses of approximately 100 randomly selected IT neurons, using few spikes in a short time window (little more than 10 milliseconds [ms]). A great deal of distributed information was found in randomly selected neurons, suggesting that almost every neuron could contribute, and it is unclear how many neurons in each area actually contribute to a given decision about a face.

Face-Selective Areas in Humans

In the human, although a few intracranial recording studies have been conducted in patients who will eventually undergo surgery for intractable epilepsy, most of the evidence for face-selectivity comes from less invasive methods, such as fMRI, positron emission tomography (PET), electroencephalogram (EEG), and magnetoencephalography (MEG). PET studies by Justine Sergent and colleagues first identified a region in the fusiform gyrus involved in face identification, which was confirmed and further investigated using fMRI in studies by Nancy Kanwisher and colleagues, under the name fusiform face area or FFA. The FFA in the right hemisphere is the most studied face-selective region of the human brain, likely because it shows robust selectivity and is found in most people in a similar location. However, this functional area cannot be localized anatomically without the use of what is called a "functional localizer" comparison between the response to faces and the response to other objects. A left FFA is also observed, typically a little smaller and more variable in its location than in the right hemisphere; an occipital fusiform area (OFA) can also be found in both hemispheres, as well as face-selective regions in the superior temporal sulcus and in the anterior temporal lobe (especially in PET because fMRI picks up a limited signal in this part of the brain). Similar to work in the monkey, human studies suggest the OFA and FFA are more important for identification, whereas face-selective regions of the superior temporal sulcus (STS) appear more important for processing eye gaze, expression, and dynamic facial movements. The homologies between the different human face-selective areas and the various face patches observed with fMRI in the monkey brain have yet to be determined, but the recent use of the same method in both human and nonhuman primates offers some of the best hope for unraveling the functional anatomy of face perception.

One controversy in the interpretation of hot spots of selectivity for categories, such as the FFA, stems from experiments demonstrating a great deal of distributed and overlapping responses to objects from various categories in the visual system, as well as the rest of the brain. One view, supported by recent work with classifier algorithms, is that faces are represented in a distributed manner just like any other category, without a unique status for the FFA. Another view is that the recognition of most non-face objects is supported by distributed representations but that faces are different and more dependent on a localized representation. There is no consensus at the moment on this question, although recent results using monkey fMRI suggest that face-selective patches form a tightly interconnected network relatively isolated from surrounding areas.

Event-related responses measured in EEG (or MEG) in humans reveal a face-selective potential that peaks at 170 ms post-stimulus onset (it is conventionally called the N170) that is typically best measured at occipitotemporal electrodes and appears to have a source near the face-selective areas in the ventral cortex. The N170 is slightly but robustly delayed by the inversion of faces and it is highly dependent on the presence of the eyes in the face. Recently, a similar potential has been measured in scalp recordings in the monkey, which offers further promise of bridging animal and human work that traditionally depended on different methods.

Face-Selective Deficits

If systems are specialized for face processing and they are to some extent localized in the brain, then it should be possible for brain damage to sometimes result in lesions that selectively impair face recognition. Such deficits, called prosopagnosia, are rare and generally result from damage to the occipitotemporal cortex. The lesions that result in prosopagnosia are varied in their specific location, and the patients vary greatly in the exact nature of their impairment. A similar deficit that does not result from acquired brain damage (sometimes called congenital or developmental prosopagnosia) has recently been associated with abnormal anatomy of the anterior fusiform gyrus.

In addition, the study of individuals with autism spectrum disorders (ASD) has revealed reduced activation of the visual system by face stimuli, in line with behavioral difficulties in face processing. Individuals with an ASD present with multiple difficulties (e.g., linguistic and communicative impairments), but it is thought that deficits in social functions may be related to abnormal face processing. Recent work has revealed fewer and smaller neurons in the fusiform gyrus of individuals with autism. Whether the functional and anatomical abnormalities cause the face recognition deficit or are a result of social impairments that limit experience with faces is unclear.

Controversies About the Nature of Processes and Origins

In addition to the controversy regarding whether the representation of faces is highly localized or distributed, two other debated issues relevant to face-selective neural responses have to do with the origins of the selectivity and the domain-specificity of the system.

What are the origins of face-selective responses? One option is that face selectivity is innately specified because of the importance of recognizing faces for any social species. Finding evidence to directly support the innateness view is difficult in principle. For instance, there appears to be a congenital aspect to developmental proposagnosia, but it has been proposed that the deficit impaired in these individuals has to do with configural processing more generally. Most of the evidence consistent with an innate face processing mechanism would come from an absence of evidence that learning is involved, but quite a bit of evidence suggests changes associated with development and more directly learning. fMRI in children show that the FFA is not present until adolescence. Work in adults shows that faces from less familiar races elicit less activity in the fusiform gyrus. Human studies of perceptual experts in non-face object domains (e.g., cars, birds, or new objects taught in the laboratory) have revealed behavioral effects (such as the inversion effect) and neural effects (such as increases in the N170 and FFA activity) that are generally thought to be the hallmark of face processing. Neurophysiology also reveals increased selectivity for trained objects in the monkey and sheep. The nature of the overlap in cells that respond to faces and objects in experts is unknown at the moment, but what is clear is that learning can lead to selectivity of the kind that is observed for faces, close to the areas that are face-selective, providing an hypothesis for the origins of face-selectivity.

A second debate does not concern the origins of the face-selective mechanisms, but whether they can only be engaged by faces (their domain-specificity). That is, regardless of whether it has been learned or is innate, there may be a system that is uniquely engaged by faces, even if closely adjacent to neural networks performing similar computations on non-face objects. Some of the strongest claims having to do with the domain-specificity of the face module come from the study of brain-damaged patients. However, this issue avoids the question of how to define what a "face" is, without which it is impossible to determine if a system is face-specific. Some of the evidence bearing on this question comes from behavioral studies of competition between face processing and other domains of expertise, suggesting a common bottleneck, which has been associated with the face-selective N170 potential. Evidence of such competition suggests that face processing is not fully functionally independent from object perception, at least not in the case of objects of expertise that are processed using a similar strategy to that of faces. Where in the brain this competition arises from (for instance at the level of the FFA or not) is not known.

Isabel Gauthier

See also Agnosia: Visual; Experience-Dependent Plasticity; Feature Integration Theory; Modularity; Neural Representation/Coding; Perceptual Expertise

Further Readings

Allison, T., Puce A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex, 9*(5), 415–430.

Behrmann, M., Avidan, G., Gao, F., & Black, S. (2007). Structural imaging reveals anatomical alterations in inferotemporal cortex in congenital prosopagnosia. *Cerebral Cortex, 17*(10), 2354–2363.

Gauthier, I., Skudlarski, P., Gore, J. C., &Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191–197.

Gross, C. G. (1992). Representation of visual stimuli in inferior temporal cortex. *Philosophical transactions of the Royal Society of London Series B, Biological Sciences, 335*(1273), 3–10.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302–4311.

Rossion, B., Kung, C. C., & Tarr, M. J. (2004). Visual expertise with nonface objects leads to competition with the early perceptual processing of faces in the human occipitotemporal cortex. *Proceedings of the National Academy of Sciences USA, 101*(40), 14521–14526.

Feature Integration Theory

Despite the apparently indisputable observation that we perceive a world of coherent objects in which the visual characteristics of each object are bound together, feature integration theory (FIT) is based on a rather contrary set of assumptions. The original theory was proposed by Anne Treisman and Garry Gelade in 1980 and was based on the view that the different visual characteristics of a given object are treated independently of each other within the brain. Accordingly, the human visual system invokes a complex divide-and-conquer set of operations whereby the different visual characteristics are dealt with by different neural systems. The visual properties of objects are defined relative to corresponding visual dimensions. For example, an object's size, color, shape, and motion (i.e., the visual features) are specified as values on corresponding dimensions, such as a large, yellow taxi moving away. Within the theory, the color, shape, size, and motion of the taxi are initially processed via separate and dimension-specific neural mechanisms. Each visual dimension is associated with a particular neural system.

Having made claims about the initial operation of the separable systems that exist to detect different visual features, the theory also confronted the problem of how the impression of coherent objects comes about. According to the theory, subsequent to the initial stage of feature registration, there is a stage of feature integration. Once the different features have been registered, operations now take place to coordinate this information and integrate the features into a representation that codes a coherent object. This entry describes the neuropsychological evidence for feature integration theory, the binding problem, and theories of feature integration, as well as alternative perspectives.

Neuropsychological Evidence

What sort of evidence has been marshaled to support this view? The answer is many and varied, but here the focus is with data from neuropsychology. Consider, for instance, the variety of different forms of color blindness. Although most cases can be attributed to problems in the eyes, with cerebral achromatopsia the problems are caused by damage or injury to particular areas of the visual cortex. In such cases, the world appears to be drained of color—there is a selective deficit in processing color—but other visual abilities such as perceiving shapes, depth, and motion remain intact.

More startling perhaps and much more rare are individuals who have problems in being able to see motion following damage to the brain. Josef Zihl and colleagues described the most famous case, labeling the condition *motion blindness*. The unfortunate individual suffered a stroke that resulted in bilateral damage to the posterior parietotemporal and occipital regions of her cortex. She was described as inhabiting a kind of stroboscopic stop/start world in which things were never seen in motion but were simply here and there. Less severe cases also exhibit selective problems with motion perception even though the ability to perceive other visual characteristics remains intact. Such evidence suggests specific brain mechanisms dedicated to perceiving color and motion, respectively.

A different and fascinating case undermines the notion that analyses of shape and color are inextricably interwoven. When presented with three vertical bars colored red, white, red, respectively, or black, red, black, respectively, the individual reproduced these displays by drawing a large red rectangular blob. In other words, this person experienced great difficulties in linking a red filled-in region with its bounded contour. His perception of these sorts of pictures reflected the fact that the color seeped out from the corresponding red areas and did not honor the areas' actual boundaries. This counterintuitive finding suggests that the brain invokes independent analyzers for color and shape, respectively, despite the compelling impression of solid objects with associated colored parts. This interesting case was documented before the recent advances in brain imaging; hence, it was impossible to know the exact cause of the problems. However, tests ruled out any kind of retinal deficit, and there was no evidence of acquired brain damage. The problems seemed to reflect the development of a brain in which the cortical regions responsible for color processing were disordered.

Binding Problem

So if we accept the assumption that such visual characteristics are processed via independent and specific neural mechanisms, and that we typically see the world as inhabited by coherent objects, then the question arises as to what is responsible for coordinating the different features so that the world appears coherent. How is it that the independent visual features are bound together so that the impression of coherent objects is formed? How does the brain solve what is known as *the binding problem?* FIT provides various answers.

In its earliest incarnation, the theory was based on the idea that, at an early stage of processing, different visual features are processed by dedicated (neural) dimensional processors or *feature maps*. There is a color map, a shape map, a size map, and so on. The various features are immediately registered on their respective maps. This process of featural pick-up is outside subjective control: It is carried out pre-attentively and in parallel. So various shapes, colors, and sizes are registered across the visual field concurrently and immediately. To make sense of it all, the information registered on the various maps needs to be coordinated. Introspection suggests that we can prioritize what it is that we want to focus on, so a primary factor invoked by the theory is *selective attention*. Following feature registration, *feature binding* takes place. Attention is primarily responsible for this binding in coordinating the featural information across the different maps.

Original Theory

In the original theory, the manner in which the different visual features are combined, conjoined or integrated, can result from any one of three factors. The first is selective attention, but a particular notion is implicated in a metaphor that invokes the *mental spotlight of attention.* As we scan the world, the impression inside the focus of attention is of perfect clarity, but the details outside this "spotlight" tend to be less clear. According to the theory, features within the spotlight of attention are typically conjoined appropriately. Second, and independently of the focus of attention, the different features may be combined as a consequence of their fitting with well-worn expectations of the world—the taxi shape must be yellow. Third, and more interesting, perhaps, is the idea that features may be conjoined at random. In this context, *illusory conjunctions* are discussed. The basic idea is that, in cases where attention is overloaded or misdirected, features from different maps can be combined incorrectly. This was originally shown in the following way.

A given display contained a single row of colored alphanumeric characters—two outer-most black digits and three inner colored letters. The characters were presented briefly and curtailed by the presentation of a random pattern of black and white squares. The task was first to report the digit and then report the colored letters. It was found that participants reported incorrect combinations of colors and letters on nearly 40% of trials. So even though a red X and blue T are presented, a person reported seeing a red T or a blue X. These are illusory conjunctions.

Critical support for the theory has also arisen from *visual search experiments.* Central here is the idea of a visual display in which a number of discrete items, such as randomly positioned colored letters, is contained. The person is instructed, under what are known as reaction time conditions, to respond as quickly and as accurately as possible whether or not the display contains a predefined target item. In feature conditions, the target is a letter of a particular color or shape (e.g., a blue letter or a "C"). In feature conjunction conditions, the person has to search for a particular kind of colored letter, such as a red X. In both cases appropriate non-target letters could be red Ts and green Xs. However, in the conjunction search conditions, neither the shape nor color of the conjunction target is unique. The combination of a specific color and a specific shape is distinctive, so the person must find the correct combination of a color and a shape (i.e., a red X).

Aside from the feature versus conjunction target distinction, a critical variable is the number of letters contained within the display. Over different trials, this number is varied, typically within the range of 1 to 30 items. There is, however, only one target present in target present displays. Performance in the separate feature and conjunction search tasks is measured as a function of whether the target is present or absent and as a function of the number of items in the displays.

Evidence consistent with the theory's predictions was forthcoming from various visual search tasks. For cases where a feature target in the feature condition was present, the time to respond did not increase with the number of display items. No matter how many green and red non-target letters were present, a blue letter stood out and was responded to relatively quickly. Such a pattern of responding is known as *featural pop-out*—a distinctive feature pops out from its background—a result that accords well with the initial stage of automatic attention-free registration of visual features. In contrast, responses in the conjunction search conditions were dramatically different. Now response times increased linearly with increases in the number of items to be searched through. So whereas feature detection was taken to be an effortless and an attention-free process, conjunction detection was taken to be effortful and attention-demanding.

Performance in the conjunction search task was interpreted in terms of the deployment of the mental spotlight of attention in what is labeled a *serial self-terminating search process.* Choose an item to focus on and check off whether it fits the description of the current target (Is it a red X?). If No, then move the spotlight to the next item. If Yes, then respond "Present." Continue this one-item-at-a-time process until either the target is found or all the letters have been checked off. If a target is not found, respond "Absent." The critical theoretical point is that the identification of a particular feature conjunction comes about because of the integrating operations that are associated with the serial deployment of the mental spotlight of attention.

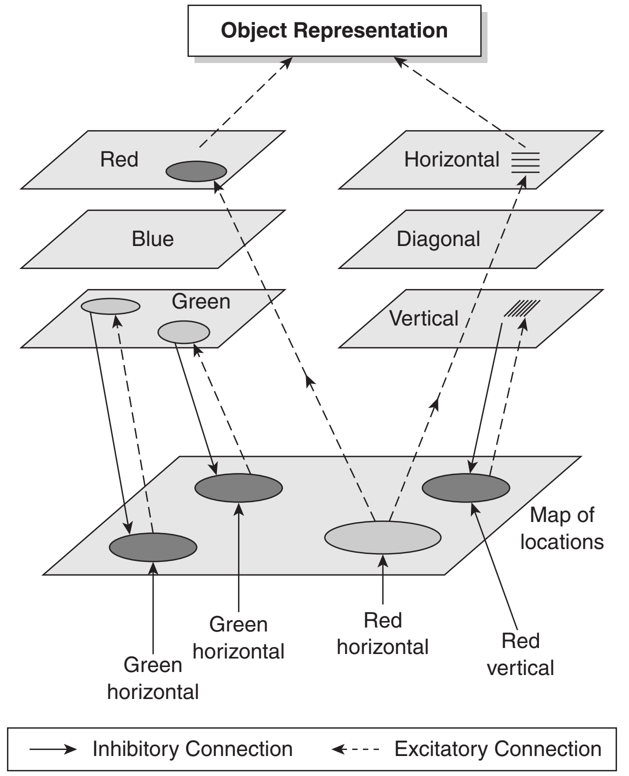


Figure 1 A Schematic Representation of Version 2 of FIT

Source: Reprinted with permission, from Treisman, A. (1999). Feature binding, attention and object perception. In G. W. Humphreys, J. Duncan & A. Treisman. (Eds.), *Attention, space and action. Studies in cognitive neuroscience* (pp. 91–111). Oxford, UK: Oxford University Press. Original Figure 6.6, p. 100.

Notes: Visual input from the world is initially registered (at the bottom) in the map of locations. Subsequently, more detailed featural analyses are carried out on the various feature maps. In this case, maps for colors and various orientations are shown. Finally, the notion of an object is captured at the level of the object representation, or, alternatively, object file.

Attention is needed to bind the separate features together. Once the features are bound together, then the impression of a coherent object comprising those bound features is generated.

Version 2

Initially, FIT provided a relatively comprehensive account of the findings from various kinds of visual search tasks. The theory appears to have galvanized the field. Much of the work on human visual attention in the ensuing years has been either directly or inadvertently influenced by FIT. Probably because of the immense interest generated by the theory, cracks soon began to appear. Cases were uncovered where (a) single visual features were rather difficult to find, and where (b) feature conjunctions were relatively easy to find. For example, when search elements were defined in terms of depth and color, then corresponding depth/color combinations were relatively easy to detect. As a consequence, version 2 of the theory was put forward (see Figure 1).

Some of the basic components were retained, but critical modifications were also introduced. For instance, now there is an "object representation" component that is also known as an *object file*. This is a temporary form of storage in which a record is made of the features that go together to define a viewed object. Once feature registration and feature integration have taken place, information regarding a particular coherent bundle of features is captured in a corresponding object file. Impressions of constituent objects in the world are based on the information contained in these object files. Such files are used to explain how a person keeps track of the objects that are currently present in a dynamically changing scene. As the objects move or change, the files are updated accordingly.

The theory is much more complex than the original, and an additional component, known as the *master map of locations,* was introduced. This map is located at an early stage of processing and captures information regarding both the position and nature of bundles of conjoined features. More detailed featural analysis now takes place following the registration of information on the master map. By allowing the master map to capture combined feature information, the theory now accommodates the fact that certain feature conjunctions are easy to find.

Alternative Perspectives

Given the complex nature of human vision, it is perhaps not so surprising that the original and simple version of FIT proved to be inadequate. Dissatisfaction with the general FIT framework for thinking has led some to abandon the theory for alternative perspectives. One such alternative, described by John Duncan, has come to be known as *biased competition.* By this view, there are multiple brain regions, each of which codes for a different perceptual feature, and these compete with one another for attentional resources. Processing is biased toward those regions that reflect most the current task requirements, and regions that win this competition reflect those aspects of the world that are currently being attended to. Nevertheless, what remains outstanding is a satisfactory account of the binding problem. FIT provides some indications of how the binding problem is solved—primarily through the application of focused attention—but a generally acceptable solution remains beyond our grasp.

Philip Quinlan

See also Attention: Physiological; Binding Problem; Psychophysics: Detection; Theoretical Approaches; Visual Search

Further Readings

Duncan, J. (2006). Brain mechanisms and attention. *Quarterly Journal of Experimental Psychology, 59,* 2–27.

Hendricks, M. I., Holliday, I. E., & Ruddock, K. H. (1981). A new class of visual defect. Spreading inhibition elicited by chromatic light stimuli. *Brain, 104,* 813–840.

Quinlan, P. T. (2003). Visual feature integration theory: Past, present and future. *Psychological Bulletin, 129,* 643–673.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Treisman, A. M., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14,* 107–141.

Zihl, J., von Cramon, D., & Mai, N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain, 106,* 313–340.

FEEDBACK PATHWAYS

The brain's pattern of connectivity involves two-way pathways: feedforward pathways and feedback pathways. For instance, if area A sends signals to area B (feedforward pathway), then area B almost invariably sends signals back to area A (feedback pathway). This entry focuses on the structural and functional aspects of the feedback pathways.

Structural Aspects

The brain's architecture reveals a modular structure, with each module performing a specialized function. In the visual system, signals travel rapidly from the retina along feedforward pathways to the input region of the visual cortex, known as the *primary visual cortex*. Signals are then processed in separate modules, each optimally tuned to a different stimulus attribute such as orientation, color, and motion. The feedforward pathways are matched by corresponding feedback pathways back to the primary visual cortex.

This pattern of two-way pathways is ubiquitous in the brain. As well as being ubiquitous, feedback pathways are massive. For example, in the pathways that link the primary visual cortex with lateral geniculate nucleus (LGN, a subcortical structure), descending fibers outnumber ascending fibers by at least a factor of 10.

A notable aspect of this pattern of connectivity is that the descending fibers do not merely complete a feedback loop, from lower-level neuron to higher-level neuron and back, but are distributed widely on their return, including many neurons in the spaces between the source neurons. This arrangement constitutes a powerful means by which a high-level area can influence the outputs of currently active cells in the low-level area and can selectively sensitize or inhibit other neighboring cells to modulate future incoming signals.

Functional Aspects

At least three major functions are served by feedback pathways: mediate the perception of objects, program (multiplex) neurons to perform different jobs at different times, and predict the course of environmental regularities (e.g., the trajectory of a ball flying through the air).

Perception of Objects

In the current view, the brain is seen as a repository of memories in the form of neural networks established early in life. This memory system contains a model of the world. Through iterative feedback signaling, it identifies objects and events.

A visual example will illustrate the concept of iterative feedback signaling. The incoming signals are first encoded in specialized modules in low-level areas of the brain and then proceed to higher areas where they activate multiple memory representations. The representations activated at the higher level are sent back via descending pathways to the primary visual cortex, which acts as a blackboard for the rest of the visual system.

Comparisons between the descending representations and the ongoing activity in the primary cortex are carried out in parallel, with each representation attempting to match itself to the pattern of activity at the lower level. This comparison involves a kind of voting as to the best fit between the ongoing activity at the lower level and the feedback representations, with the outcome sent back to the higher area. Several iterations of this kind take place, culminating in the confirmation of the memory representation that best fits the input data.

Multiplexing of Function

Feedback signals can reconfigure the same neurons to perform different functions in successive phases of processing. For example, suppose that we flash on a screen a simple visual scene containing a central figure (e.g., a square patch) on a uniformly textured background. Until about 80 milliseconds from stimulus onset, a given neuron in the primary visual cortex responds only to local features (e.g., line orientation) presented within the small part of the visual field from which that particular neuron receives its input. At this stage, the neuron is entirely "unaware" of all other elements in the display.

About 40 milliseconds later, the same neuron begins to respond to the boundaries of the central figure. Some 40 milliseconds after that, the neuron responds optimally only to the figure itself, not to the background. Thus, in less than 200 milliseconds, that neuron has served three different functions: feature detector, contour detector, and figure-ground detector.

This multiplexing of function depends critically on feedback pathways. If feedback signals from higher brain regions are prevented from reaching the neurons at the lower level (by surgical removal or cooling of the higher regions), only the initial function (feature detection) remains. By enabling a few multiplexed neurons to do the job of many single-function neurons, this functional organization of feedback pathways permits a leaner and more parsimonious system than one with enough neurons to do the same job in a feedforward fashion.

Environmental Regularities

An example of how expectations based on physical laws and environmental regularities are mediated by feedback signals is provided by studies that monitored the activity along the two-way pathways between the primary visual cortex and the lower-level LGN in the cat in response to a moving stimulus. Once the direction of motion has been computed at the higher level, signals are sent to the lower level.

The descending fibers form a feedback loop between neurons in the visual cortex and LGN neurons located next to the initial source-cell in the direction of motion. This circuitry provides an anticipatory coordinating influence by increasing the sensitivity of the LGN cells that are about to be activated by the moving stimulus.

The main point to be gathered from this pattern of two-way communications is that feedback from the primary visual cortex to LGN can prepare the system for a stimulus that is about to occur at a location in the expected path of motion. At a more general level, this is an example of how descending signals can mediate expectations based on physical laws and environmental regularities.

Vincent Di Lollo

See also Attention: Physiological; Consciousness; Cortical Organization; Object Perception; Visual Processing: Extrastriate Cortex; Visual Processing: Primary Visual Cortex

Further Readings

Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews, 36,* 96–107.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129,* 481–507.

Hawkins, J., (2004). *On intelligence.* New York: Holt.

FILM (CINEMA) PERCEPTION

Film (cinema) perception refers to the sensory and cognitive processes employed when viewing scenes, events, and narratives presented in edited moving images. Dynamic visual media such as film and television have increasingly become an integral part of our everyday lives. Understanding how our perceptual system deals with the differences between these mediated visual experiences and the real world helps us understand how perception works in both situations. This entry focuses on three of the many differences between film and reality:

- Film creates the illusion of motion through the rapid presentation of still images.
- Film creates the illusion of continuity across a cut.
- Film represents scenes and events across edited sequences of shots filmed at different places and times.

Although this list is not exhaustive, these three differences are critical for understanding how we perceive film. This entry provides a brief overview of these differences and current theories about how they are dealt with by our perceptual system.

Moving Pictures

Movies consist of a series of still images, known as *frames,* projected onto a screen at a rate of 24 frames per second. Even though the frames are stationary on the screen and are momentarily blanked as a new frame replaces the old, we experience film as a continuous image containing real motion. The two perceptual phenomena contributing to this experience are *persistence of vision* and *apparent motion.*

Persistence of vision refers to the continued perception of light after the stimulus light has been turned off. During film projection, the light is obscured by the closing of a shutter as the film moves from one frame to the next. This creates an alternation between light (shutter open—frame projected) and darkness (shutter closed) 24 times per second. Persistence of vision "fills in" the dark interval, but only partially, because a shutter rate of 24 frames per second results in a noticeable flicker. Early film used shutter rates between 12 and 24 frames per second earning them the nickname "The Flicks." Modern film projectors eradicate this flicker by blanking each frame three times, which increases the flicker rate above the *critical flicker fusion* rate of 60 hertz (Hz) and ensuring that the perception of light is continuous because of persistence of vision.

The motion we perceive in film is apparent because it is based on static visual information rather than real motion. Apparent motions can be broadly classified as *long-range* and *short-range* according to the conditions under which they are perceived. Long-range apparent motions, such as *beta movement,* are perceived when two objects are alternately presented at two different locations about 10 times a second. The two objects are perceived as a single object moving smoothly between the two locations. Because of the slow rate of presentation and the large distances covered by the apparent motion, long-range apparent motions are thought to be processed late in the visual system and require inferences based on knowledge of real motion and the most likely correspondences between objects in the image sequence.

Short-range motions occur when static images depicting only slight differences in object location are presented rapidly (> 13 Hz). Short-range motion processing occurs early in our visual system, does not require perceptual inferences to understand the motion, and is the same system used to perceive real motion. It is commonly believed that the apparent motion perceived in films is beta movement. However, although beta movement and other long-range motion phenomena such as apparent rotations and transformations may occur during film perception, they cannot account for most motion perceived in film. The 24-Hz presentation rate used in film is too fast for long-range motion and film frames are too complex, making the task of identifying corresponding objects in subsequent frames too difficult. Instead, apparent motion in film results from the same short-range motion system used to detect real motion. Motion detectors in the early visual system respond in the same way to the retinal stimulation caused by real motion and by rapidly presented (> 13 Hz) static images that depict only slight differences in object location.

This results in a sensory experience of film that is indiscernible from reality.

Editing and the Illusion of Continuity

In film, we perceive scenes and events as continuous even though they are presented across multiple viewpoints that change instantaneously across edits. This illusion is referred to as *continuity*. The mismatch between the psychologically perceived continuity and the spatiotemporally discontinuous nature of the visual information was first noted by the psychologist Hugo Münsterberg in 1916. Münsterberg hypothesized that some of these violations are acceptable because a cut away to a different viewpoint within a scene mirrors the attentional shift a viewer would naturally perform when observing the same scene in the real world.

If an edit is to function as an analogue for an attentional shift, viewers need to be able to anticipate the shift in viewpoint to update their mental representation of the depicted scene. Such a *constructivist* account of film was advocated by Julian Hochberg and Virginia Brooks, who suggested that questions arising from the events depicted in the previous shot motivate a cut to a shot that answers the question and allows the viewer to link the two shots conceptually. For example, when filming the conversation depicted in Figure 1, a cut between shot B and C can be motivated by a sudden head-turn of the character in black, creating the perceptual enquiry, "Is he about to speak?" Recent evidence by Tim Smith and John Henderson has shown that when cuts are preceded by motion onsets, viewers orient quicker to the content of the new shot and are less aware of the editing compared with cuts without such attentional cues. Smith and Henderson named this phenomenon *edit blindness*.

Editing conventions (such as the 180º rule depicted in Figure 1) rely heavily on our natural tendency to attend to social features of dynamic visual scenes. Smith has shown that when multiple viewers are presented with videos of real-world scenes, their attention will synchronize as they attend to people and track them over time. This *attentional synchrony* enables film editors to predict where multiple viewers will attend to replicate their attentional shifts through editing. Recent neuroimaging evidence has also indicated that this synchronization may extend to how the film is processed. Uri Hasson and colleagues used neuroimaging to record the brain activation of multiple viewers while watching feature films. The researchers observed a high degree of synchronization in brain regions responsible for such processes as language comprehension, emotion, and face perception. Although such synchronization does not necessarily indicate that we are all experiencing a film in the same way, it supports the idea that there is a high degree of consistency in the perceptual enquiries we employ when processing dynamic visual scenes.

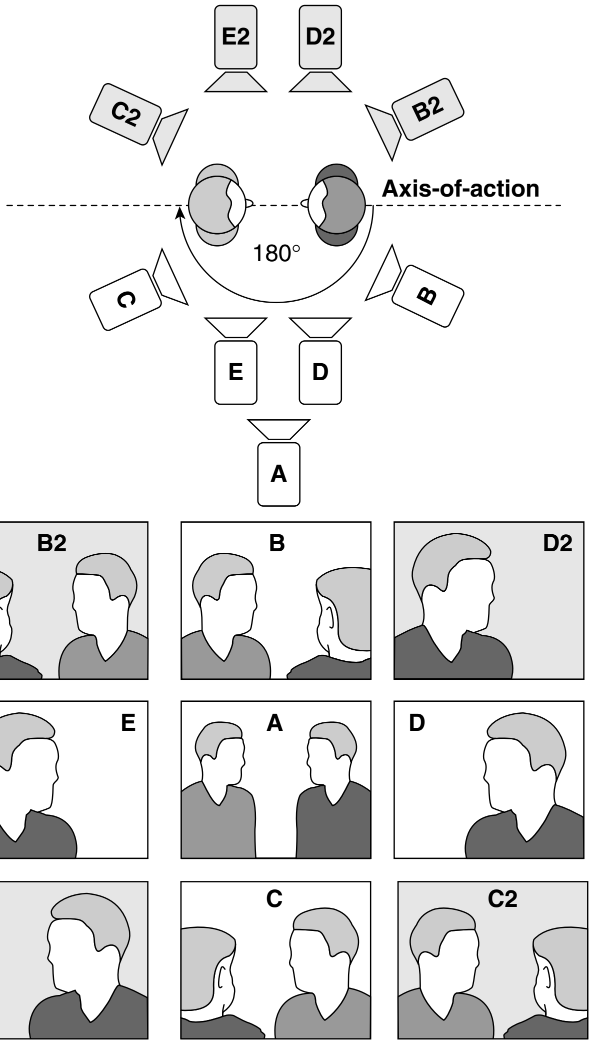


Figure 1 The 180° Editing Rule

Notes: Once the space of the scene has been established by camera A all other shots must be taken from the same side of the axis of action. A cut across the line (cameras B2, C2, D2, or E2) would create a discontinuity.

The Perceptual Construction of Space

One of the assumed benefits of adhering to classical editing conventions is that they aid viewer comprehension of the depicted space. For example, the 180° rule states that when filming a scene, all shots should be filmed from the same side of the *axis of action,* for example, the line joining the characters involved in the conversation. Any sequence that crosses the line (e.g., a cut from a white to a gray camera shown in Figure 1) is believed to confuse the viewer and lead to disorientation. This hypothesis has received support from an experiment conducted by Uta Frith and Jocelyn Robson. Children were presented with simple films that either adhered to or broke the 180° rule. Frith and Robson found that the children who saw the conventional version were able to reconstruct the film more accurately than were children shown the unconventional version. Hochberg and Brooks explained such an effect as evidence of the viewer's inability to construct a coherent spatial representation of the scene because crossing the line (a) removes landmarks such as background features that would normally be used to identify the relationship between shots and (b) violates viewer expectations about the location and direction of objects on the screen such as the left-right relationship of the conversational partners in Figure 1.

However, recent evidence from Dan Levin and Dan Simons has questioned the degree to which we attend to, encode, and monitor details within a film. Participants were shown a video depicting two people having a conversation. Every time a cut to a new shot occurred, at least one continuity error was inserted such as the disappearance of a scarf. When participants first watched the film without being told that there might be a change, 90% failed to spot any continuity errors. In another video depicting a woman getting up from a desk in the first shot and answering a phone in the second, 66% of viewers failed to notice that the actress changed across the cut. Levin and Simons interpreted these findings as indicating that, rather than maintaining a highly detailed and coherent representation of the depicted scene, viewers only encode a small amount of the depicted visual information. Instead, continuity of space and time is assumed and perceptual inferences about the location and form of minimal details such as people and their movements are monitored to test the validity of the continuity assumption. If this inferred continuity hypothesis is true, the role of the editing conventions may be to facilitate these perceptual inferences and ensure that they can be satisfied following a cut.

One important condition for continuity seems to be the location of objects on the screen. Géry d'Ydewalle and colleagues recorded viewer eye movements while they watched videos edited with or without violations of the 180° rule. The researchers observed a peak in saccadic eye movements (i.e., attentional search) following cuts that violated the rule and concluded that viewers anticipate the screen location of objects and when these expectations are violated (e.g., a cut from B to C2 in Figure 1), viewers have to repair their representation, leading to a break in perceived continuity. The critical nature of attentional shifts across cuts for the inference of continuity was formalized by Smith as the *attentional theory of continuity editing*. Viewer attention throughout a film specifies which visual features are represented in memory, how perceptual enquiries are formulated and tested, and whether continuity can be inferred from the satisfaction of minimal expectations across cuts.

Modern cinema and television are an integral part of our everyday lives. Except for a few theoretical and empirical pioneers, however, the big questions of film perception have received little psychological attention. With new psychological methods such as eye tracking and neuroimaging at our disposal, perhaps now is the time for Münsterberg's 1916 declaration of film as the domain of the psychologist to finally come true.

Tim J. Smith

See also Attention: Object-Based; Attention: Spatial; Change Detection; Depth Perception in Pictures/Film; Event Perception; Motion Perception; Pictorial Depiction and Perception

Further Readings

Frith, U., & Robson, J. E. (1975). Perceiving the language of films. *Perception, 4*(1), 97–103.

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D. J. (2008). Neurocinematics: The neuroscience of film. *Projections, 2,* 1–26.

Hochberg, J., & Brooks, V. (1978). The perception of motion pictures. In E. C. Carterette & M. P. Friedman

(Eds.), *Handbook of perception* (Vol. 10, pp. 259–304). New York: Academic Press.

Levin, D. T., & Simons, D. J. (2000). Perceiving stability in a changing world: Combining shots and integrating views in motion pictures and the real world. *Media Psychology, 2,* 357–380.

Münsterberg, H. (1916). *The photoplay: A psychological study.* New York: D. Appleton.

Smith, T. J. (2006). *An attentional theory of continuity editing.* Doctoral thesis. University of Edinburgh, Edinburgh, Scotland. Retrieved from http://homepages.inf.ed.ac.uk/s9732397/publications/smith_ATOCE_2006.pdf

Smith, T. J., & Henderson, J. M. (2008). Edit blindness: The relationship between attention and global change blindness in dynamic scenes. *Journal of Eye Movement Research 2,* 1–17.

d'Ydewalle, G., Desmet, G., & Van Rensbergen, J. (1998). Film perception: The processing of film cuts. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 357–367). Oxford, UK: Elsevier.

FLAVOR

All foods and beverages stimulate multiple, physiologically distinct, sensory systems. Information from volatile chemicals reaches the olfactory receptors deep inside the nose either via sniffing (*orthonasally*), or via the nasopharynx once the food is in the mouth (*retronasally*). With either route, information reaches the olfactory bulb via cranial nerve (CN) I. Molecules from soluble compounds in foods—including (but not limited to) carbohydrates, salts, acids, amino acids, and probably fatty acids—bind with taste receptors, most of which are embedded within raised structures (papillae) on the dorsal and lateral surfaces of the tongue. For taste information, these structures are innervated primarily by CN VII (chorda tympani) and CN IX (glossopharyngeal). Foods also stimulate nerve fibers responsive to touch and temperature, and—in the case of hot spices—those that mediate pain, particularly via branches of CN V (trigeminal) located throughout the mouth, tongue, and nose. Unique receptor mechanisms and specific nerve pathways from the periphery to discrete areas at each level of the brain distinguish these sensory modalities, in the same way that vision and hearing are distinct from one another.

In everyday considerations of food or beverage flavors, the word *taste* is often used, as in an *orange taste*. However, this usage not only covers actual tastes, typically sweetness and sourness, but implicitly also the olfactory quality, orange, as it is perceived retronasally. Given the multisensory nature of eating and drinking experiences, flavor is most properly seen as an aggregation of odor, taste, tactile, and, under some circumstances, visual properties. However, these apparently distinct sensory modalities are seldom independent. What we respond to, perceptually and hedonically, during food consumption is not a collection of discrete sensory signals, but rather an overall percept of a flavor. Thus, orange juice flavor appears somehow more than the sum of its parts: the odor of orange, the tastes of sugar and citric acid, and a variety of tactile and temperature sensations. This is rather in contrast to multisensory interactions involving other sensory systems. For example, we automatically link the visual image of a bird and its song because they arise from the same location, but we are nevertheless conscious that these aspects are distinct in sensory terms. In contrast, flavor components seem more highly integrated. As such, the study of flavor perception is interesting both as an example of commonplace multisensory interaction and integration and for what this integration of the different components of flavor—and in particular, odors and tastes—reveals about the biology and psychology of food perceptions.

One implication of the interchangeability of *flavor* and *taste* in common usage is that we routinely fail to make a distinction between olfactory and taste qualities within flavors. In turn, this is one manifestation of the fact that, during normal food consumption, these flavor elements are highly integrated. In studies of other sensory systems—vision and audition, for example—such cross-modal integration is inferred from the influence of one modality on another. Commonly, this is an enhanced (sometimes supra-additive) response to information from one sensory system resulting from concurrent input from another modality. For example, in a crowded room, speech comprehension is improved if we see the speaker's lip movements. There is similarly evidence that tastes and odors, when encoded together as a flavor, interact to modify one another. This entry discusses various aspects of flavor, focusing on multisensory, cross-modal, and attentional processes and neural representation.

Multisensory Integration and Learned Sensory Congruency

The most obvious expression of odor-taste interactions is the widely observed attribution of taste qualities to odors. When asked to describe the odor of caramel or vanilla, most people will use the term *sweet-smelling;* similarly, *sour* is used for the odor of vinegar. This commonplace phenomenon could be dismissed as merely imprecise language (because highly specific odor descriptors are elusive) or even metaphor, given that the odor name is likely to refer to an object, which might also be sweet or sour. However, there are measurable consequences of odor taste qualities, in that these odors, when added to tastants in solution, can modify the taste intensity. The most common finding relates to the ability of food odors such as strawberry or vanilla to enhance the sweetness of sucrose solutions. This phenomenon is both taste and odor specific. For example, the sweet-smelling odor of strawberry will enhance a sweet taste, but the odor of bacon will not. Conversely, a nonsweet taste, such as saltiness, will not be enhanced by strawberry. Evidence for these interactions is apparent even at subthreshold levels, when a concurrent taste can substantially increase the detectability of the odor, but only when the taste and odor share a common property, such as (tasted and smelled) sweetness. Behaviorally, too, odor-taste interactions are evident. Thus, the presence of a sweet- or sour-smelling odor enhances response times to name sweet and sour tastes, respectively.

Studies of interactions between visual, auditory, and somatosensory systems have demonstrated the importance of spatial or temporal contiguity in facilitating cross-modal sensory integration. The importance of taste-related odor properties for understanding sensory integration in flavors derives principally from the fact that these effects only arise once the odor and taste have been repeatedly experienced together in spatiotemporal synchrony—that is, simultaneously in the mouth—and typically in the context of foods or beverages. As a result of such co-exposure, a process known as *configural learning* occurs. In this process, the individual elements—for example, the taste and the odor—become encoded together as parts of a compound stimulus. As a result, the perceptual boundaries of the odor and taste become less well defined as they become dimensions of this compound, the flavor, rather than being independent sensory modalities. Perceptually, the odor and taste can be seen to become congruent—in essence, belonging together—because they possess similar qualities. Hence, the sweetness of a taste such as sucrose is seen to be more congruent with the sweet-smelling odor of caramel than it is with the odor of bacon, which typically has no sweet smell.

This process of learned congruency has been repeatedly demonstrated experimentally. Novel odors that have little or no smelled sweetness or sourness take on these qualities when repeatedly paired with sucrose or citric acid, respectively, in solution. Only following this co-exposure does the odor have the capability of enhancing a (now) congruent taste. Given the associative origin of these effects in the context of foods and beverages, we might expect cross-cultural differences in the extent to which particular odors and tastes are judged as congruent. For example, the odor of pumpkin is likely to smell sweeter in those cultures where it is incorporated into desserts (e.g., the United States) compared with cultures where it is savory. Such acquired perceptual similarity has been seen as an example of a *learned synesthesia* in which qualities in one sensory system (olfaction) are able to evoke qualities in another (taste) as a result of frequent co-occurrence. The joint encoding appears automatic as this form of learning appears not to require conscious awareness by the subject of the particular odor-taste contingencies.

One explanatory model for these effects proposes that each experience of an odor always invokes a search of memory for prior encounters with that odor. If, in the initial experience of the odor, it was paired with a taste, a configural (flavor) stimulus is encoded in memory. Subsequently sniffing the odor alone will evoke the most similar odor memory—the flavor—that will include both the odor and the taste component. Thus, for example, sniffing caramel odor activates memorial representations of caramel flavors, which includes a sweet component. This results in either perceptions of smelled taste properties such as sweetness or, in the case of a mixture, a perceptual combination of the memorial odor representation with the physically present taste in solution.

Cross-Modal Binding

In vision, aspects of a scene or object include features such as form, color, and movement, and researchers have considered how these different aspects of the one scene or object form a coherent perception. The neural processing of form can be shown to be independent of that of color, but our perception is always that the two visual phenomena are bound seamlessly together. To understand flavor perception, it is also crucial to know the mechanisms responsible for binding tastes, odors, and tactile sensations into one coherent percept.

The most influential model of visual binding proposes that individual visual features are only loosely associated during early stages of processing, most likely by a common spatial location, but are bound to form a coherent perception only as a result of attention directed toward combining these features as aspects of the same object or scene. In flavors, the different stimulus elements are associated temporally. However, although both gustatory and somatosensory receptors are spatially located in the mouth, olfactory receptors are not. The question then arises of how odors become bound to taste and touch. That such binding does occur is reflected in the *olfactory location illusion,* in which the odor components of a food appear to originate in the mouth. Thus, we *never* have a sense that the "oranginess" of orange juice is being perceived within the nose, even if we are aware that it is an odor. This illusion is both strong and pervasive, even though we are frequently presented with evidence of the importance of the olfactory component in flavors, such as through a blocked nose during a head cold.

The olfactory location illusion is effectively an equivalent phenomenon to the auditory-visual ventriloquism effect in that, like the ventriloquist's voice, the location of the odor is captured by other sensory inputs. The extent to which either concurrent taste or somatosensation, or both, is chiefly responsible for the capture and referral of olfactory information to the oral cavity is not known exactly. However, the somatosensory system is more strongly implicated because, like vision, it provides detailed spatial information, which the taste system does not. Moreover, in neuroimaging studies, odors that are available to bind with tastes—that is, those presented retronasally (via the mouth)—have been shown to activate the mouth area of the primary somatosensory cortex, whereas the same odors presented via the nose do not. This distinction, which occurs even when subjects are unaware of route of stimulation, suggests a likely neural correlate of the binding process and supports the idea that somatosensory input is the underlying mechanism.

Because sniffed odors, even of the same quality, can be distinguished from odors in the mouth, there is a strong suggestion that the two routes of stimulation are processed with some independence, which is supported by neuroimaging studies that show different activation patterns in cortical olfactory areas as a result of route of administration. From an adaptive point of view, this makes sense. Olfaction has been described as the only dual sense because it functions both to detect volatile chemicals in the air (orthonasal sniffing) and to classify objects in the mouth as foods or not, and each of these roles has unique adaptive significance. Because the mouth acts as the gateway to the gut, our chemical senses can be seen as part of a defense system to protect our internal environment—once something is placed in the mouth, there is high survival value in deciding whether consumption is appropriate. Sensory qualities (tastes, retronasal odors, tactile qualities) are therefore combined into a single perception that identifies a substance as a food. This concept of flavor integration is consistent with the ecological view of perception proposed by psychologist J. J. Gibson in the 1950s and 1960s. Gibson suggested that the physiological origin of sensations was less important than that the sensations—singly or in concert—could be used to identify objects. Hence, what matters is that the *object itself* has certain qualities that can be identified. The implication of this interpretation for the perception of foods is that flavor is a functionally distinct sense that is cognitively "constructed" from the integration of distinct physiologically defined sensory systems (such as olfaction and gustation) to perceive and identify objects that are important to our survival.

Attentional Processes

There appears to be a role for attentional processes in binding the elements of flavor, just as there is with the visual features of an object. Although the simple repeated experience of an odor and sweet taste together appears to be sufficient to produce changes in the smelled sweetness of the odor, the ability of that odor to later summate with the taste to produce enhanced sweetness relies on a concurrent attentional approach that combines the odor and taste, rather than identifying them as separate elements in the flavor. In other words, for a complete binding of flavor features via configural learning, synthesis of the elements via attending to the whole flavor is critical. In attending to these elements jointly, we are locating them not just to a spatial location in the mouth but also binding them to an object—the food.

The effects of attentional manipulations show that the analogy between the cross-modal binding of qualities in flavors with the binding of visual features is limited. Hence, we can, if required, switch between a synthetic approach to flavor and an analysis of the flavor elements. So apple flavor can be both a synthetic percept and, with minimal effort, a collection of tastes (sweet, sour), textures (crisp, juicy), and odor notes (lemony, acetone-like, honey). A more precise way of conceptualizing flavor is that cross-modal sensory signals are combined to produce a perception, rather than combining synthetically—in the way that odors themselves do—to form a new sensation. The analysis of flavors into their discrete elements is often undertaken by wine tasters or panels of trained food or beverage tasters, for example. During normal food consumption, on the other hand, we typically respond to flavors synthetically—an approach reinforced by the olfactory illusion and by the extent to which flavor components are congruent. As noted earlier, this implies a sharing of perceptual qualities such as sweetness of a taste and of an odor, derived from prior experience of these qualities together.

The binding of odors, tastes, and other sensory inputs through attention is likely to have an important role in facilitating our hedonic responses to foods. Such responses tend to be unitary and immediate—a single global ("gut") accept or reject response to the overall flavor—rather than an analysis of the relative merits of the individual odor, taste, or tactile components. Such analysis may actually inhibit flavor appreciation. Although you may not have the vocabulary of an expert who analyzes a wine, you might respond to that wine in a way that enhances your enjoyment of its flavor. The binding of sensory qualities in flavors therefore intrinsically has consequences for hedonic judgments of acceptability, suggesting a means through which integration acquires its adaptive value.

Neural Representations of Flavor

Consistent with the models based on psychophysical data, recent research has provided neural evidence for a functional flavor network that engages cortical areas involved in processing of odor, taste, and somatosensory information. Human neuroimaging (predominantly functional magnetic resonance imaging [fMRI]) studies have identified key processing areas for olfaction in the piriform and orbitofrontal cortex, although odor-related responses have also been recorded in the insula and operculum, which are also primary taste processing areas. Data suggest that orthonasal and retronasal stimulation activate different areas within the flavor network, but only when the odor is identified as part of a flavor. Distinctions are evident, for example, within the orbitofrontal cortex, with orthonasal stimulation activating the caudo-lateral areas and retronasal stimulation, the medial area.

Although many of these areas are active in response to odors and tastes, more important from the point of view of understanding the neural basis of flavor perception is that *congruent* odor-taste pairs preferentially activate a network of sites. This includes areas within the anterodorsal insula/frontal operculum, caudal orbitofrontal, and anterior cingulate cortex, which receive projections from both the orbitofrontal cortex and insula. Crucially, these areas demonstrate supra-additive fMRI responses to such congruent pairs (e.g., *vanilla* + *sweet*), relative not just to the odors (*vanilla*) and tastes (*sweet*) presented individually, but also relative to other, non-congruent odor-taste pairs (e.g., *vanilla* + *salty*). Because the difference between congruent and incongruent odor-taste pairs is prior experience with the elements as a flavor, such discrete activation represents a network that responds to multisensory stimuli and is critically dependent on learning.

Recordings at the level of individual neurons provide a similar picture of specialization for flavor that is distinguishable from single sensory modalities. Cells in the caudo-lateral orbitofrontal cortex and ventral insula of the monkey have been shown to be responsive to unimodal olfactory, gustatory, somatosensory, and visual stimulation. Identified among these were multimodal neurons that responded specifically to tastes and odors that had occurred together in flavors, such as the sweetness of glucose and fruit odors, rather than to incongruent combinations such as saline and these same odors. It is thought that such multimodal neurons develop from unimodal neurons through learning of appropriate combinations of signals during repeated pairing of particular tastes and odors. Such findings once again point toward a plastic neural flavor network responsive to prior experience of the odor as part of a flavor.

Our current understanding of flavors is thus that they consist of unitary percepts constructed from repeated concurrent experience of multimodal sensory—tactile, taste, and odor—information perceived to originate from the oral cavity. In turn, such perceptions rely on attentional allocation that emphasizes binding of these sensory signals, rather than their analysis. Consistent with Gibson's idea of a functional sensory system, a cortical flavor network engages multiple areas involved in unimodal sensory processing, but only when the combination of sensory inputs has some adaptive meaning for identifying a global sensory quality that uniquely identifies foods or beverages.

J. Prescott

See also Feature Integration Theory; Multimodal Interactions: Color–Chemical; Olfaction; Olfactory Quality; Synesthesia; Taste; Taste Stimuli: Chemical and Food

Further Readings

Dalton, P., Doolittle, N., Nagata, H., & Breslin, P. A. S. (2000). The merging of the senses: integration of subthreshold taste and smell. *Nature Neuroscience, 3,* 431–432.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Prescott, J. (2004). Psychological processes in flavour perception. In A. J. Taylor & D. Roberts (Eds.), *Flavour perception* (pp. 256–277). London: Blackwell.

Rolls, E. T., & Bayliss, L. L. (1994). Gustatory, olfactory, and visual convergence within the primate orbitofrontal cortex. *Journal of Neuroscience, 14*(9), 5437–5452.

Rozin, P. (1982). "Taste-smell confusions" and the duality of the olfactory sense. *Perception and Psychophysics, 31,* 397–401.

Small, D. M. (2004). Cross-modal integration: Insights from the chemical senses. *Trends in Neuroscience, 27,* 120–123.

Small, D. M., & Prescott, J. (2005). Odor/taste integration and the perception of flavor. *Experimental Brain Research, 166,* 345–357.

Stevenson, R. J., & Boakes, R. A. (2004). Sweet and sour smells: The acquisition of taste-like qualities by odors. In G. Calvert, C. B. Spence, & B. Stein (Eds.), *Handbook of multisensory processes* (pp. 69–83). Cambridge: MIT Press.

Stevenson, R. J., Prescott, J., & Boakes, R. A. (1999). Confusing tastes and smells: How odors can influence the perception of sweet and sour tastes. *Chemical Senses, 24,* 627–635.

FRAGRANCES AND PERFUME

Fragrance is a generic term that refers to odors, smells, scents, or aromas having a positive connotation. Fragrances are used nearly ubiquitously in commercial production (e.g., "new car" smell). Even "unscented" products are scented with fragrances that mask the odors involved in production. The term *fragrance,* or more often *fine fragrance,* is also a synonym for *perfume.*

Perfume is a subtype of fragrance for the purpose of self-scenting and environmental embellishment. The central function of perfume is to create a hedonically pleasing olfactory experience. It differs from many other perceptual phenomena for being primarily produced and demanded for sensory pleasure rather than physiological necessity. Other functions of perfume are to make the wearer appear well groomed or to belong to a particular social class or demographic, and to cover up fetid body or household odors. Thus, perfume facilitates social interactions. This entry describes the history of perfume use and production, ingredients of modern perfumes, perfume qualities, perfume development, and the psychology and physiology of perfume perception.

History of Perfume Use and Production

The word *perfume* comes from the Latin *per fumum,* meaning *through smoke,* and originated about 4,000 years ago among the Mesopotamians in the form of incense. Perfumery, the art of making perfume, then traveled to Egypt where it was initially only used in rituals for the gods or pharaohs. The original aromatic essences were herbs and spices such as coriander and myrtle; flowers were not used until much later. The Persian philosopher and physician Avicenna (ca. 980–1037) introduced the process of extracting oils from flowers by *distillation;* the method of boiling a liquid mash through which chemicals with different properties can be separated. This method is still used.

Perfumery came to Europe in the 14th century when, at the behest of Queen Elizabeth of Hungary in 1307, the first modern perfume made of scented oils blended in an alcohol solution was produced. This perfume was thereafter dubbed "Hungary Water." During the Renaissance, France became the epicenter of perfume development and culture, and by the 14th century cultivation of flowers in the south of France had grown into a major industry. By the 18th century, the Grasse region of France had a booming perfume trade, and it remains the capital of the Western perfume world. "Cologne" was originally invented by Italian perfumers living in Koln (Cologne) Germany in the 1700s and was made from rosemary and citrus essences dissolved in wine. The scent became quite popular, was a favorite of Napoleon Bonaparte, and was even believed to be a preventative for the plague. Today, the term *cologne* is a generic term for a weakly concentrated perfume or a man's fine fragrance.

Ingredients of Modern Perfumes

Contemporary perfumes contain between tens to hundreds of ingredients and comprise the following: (a) essential oils derived from natural aromatic plant extracts or synthetic aromatic chemicals that are classified by structural group (e.g., alcohols, esters, aldehydes, and terpenes); (b) fixatives, natural or synthetic substances used to reduce the evaporation rate, increase perceived odor strength, and improve stability; and (c) solvents, the liquid in which the perfume oil is dissolved. The typical solvent solution is 98% ethanol and 2% water.

The complete use of natural ingredients in perfumes today is rare because of the labor and expense involved. However, a recently invented technique, called *living flower* is an effective compromise. In this technique, the specific molecules emitted by a living rose (for example) are measured using gas chromatography-mass spectroscopy. A combination of two techniques whereby an analytical chemist can both qualitatively and quantitatively evaluate a solution containing a number of chemicals. The chemical composition of the natural rose is thereby determined and then recreated synthetically. Synthetics are important perfume ingredients for various reasons including lasting power and lower allergic reactivity. Some synthetics scents do not exist in nature, which provides the perfumer an even wider palette.

Perfume Qualities

The concentration level of the perfume oil in a fine fragrance indicates its intensity and its predicted duration on the skin. The more concentrated the perfume is, the stronger the scent is, and the longer it will last. Although there is variability within the definitions, there are four major perfume classifications. *Parfum* contains between 15 and 30% aromatic compounds*; eau de parfum* contains 8 to 15% aromatic compounds; *eau de toilette* ranges from 4 to 8% aromatic compounds; and *eau de cologne* contains between 2 and 5% aromatic compounds.

The olfactory aspects of a perfume are typically described in musical metaphors because there are few words that specifically describe olfactory experience. Thus, the combination of ingredients in a perfume is called a *composition,* and it has three *notes* that unfold over time. The first note is called the *top note,* or *head note*, and it produces the immediate impression of the perfume. Top notes consist of small, light molecules with high volatility that evaporate quickly. *Middle notes* (also called *heart notes*) emerge just before the top notes have dissipated. Scents from this note class appear anywhere from 2 minutes to 1 hour after the application of a perfume. *Base notes* appear while the middle notes are fading. Compounds of this class are often the fixatives used to hold and boost the strength of the lighter top and middle notes. Base notes are large, heavy molecules that evaporate slowly and are usually not perceived until 30 minutes after the application of the perfume. Some base notes can still be detected 24 hours or more after application. The varying evaporation rates of different molecules in a perfume mean that a perfume will not smell the same when it is first put on as it does 3 hours later. How long you have owned a perfume can also affect its qualities. The biggest culprit is exposure to daylight, which contains photons of sufficient energy to break the chemical bonds in aromatic compounds. A brownish color in a perfume indicates that it has absorbed light. In dim light or darkness, a perfume can last hundreds of years.

Data from the fragrance industry, most of it unpublished, suggests that skin type may influence the quality and longevity of a perfume. In particular, youthful, moist skin is the best for aroma fidelity and duration, and oily skin is better than dry skin. The reason is that dense and hydrated skin cells hold the odor molecules on the skin surface so that the fragrance can radiate off it. Oily skin also enables the odor molecules to bind to the skin surface, but dry skin absorbs the perfume liquid most quickly and the scent is less detectable. Endogenous body odors, because of consumption of volatiles (e.g., cigarette smoke, alcohol, garlic), health and unique physiology, can combine with the fragrance an individual is wearing and alter the way it smells.

Perfume Development

The science of perfume is chemistry, the aromatic result is art. A perfume today is typically created as follows: A beauty/fragrance company will approach a large aromachemical company (which will produce the perfume, though it will be branded by the beauty company) with a *brief*. The brief is an abstract description of what the perfume should convey. For example, "a perfume that will be ideal for the successful, independent professional woman in her 30s who is environmentally conscious and seeking romance and we want this perfume to make her feel beautiful and desirable." Note that there is no mention of actual fragrance qualities. The aromachemical company then charges one of its perfumers to interpret this abstract description into a fragrance. After a considerable number of assessments, changes, and tweaking, a formulation is finally prepared and marketed.

Perfumes are not usually tested by consumers before market release; however, fragrances for household products such as detergents and shampoos often are. In fragrance testing for consumer products, panels of both experts (evaluators who work for the fragrance company) and untrained consumers assess various formulations of a given scent. As a function of the evaluations, specific scent formulations are then changed or approved.

There are no regulatory policies enforced in perfumery, and therefore ingredients do not have to be listed on packaging; a trade necessity to minimize pirating because perfumes cannot be trademarked or patented. The safety of fragrance compounds themselves, however, is monitored and must meet Generally Regarded as Safe (GRAS) and Occupational Safety and Health Administration (OSHA) standards.

The Psychology and Physiology of Perfume Perception

It is unlikely that two people will perceive a perfume in exactly the same way. The reasons for this range from experience to culture, age, and physiology. Most importantly, past experiences with a scent will result in different hedonic associations and preferences. For example, grandmother wore *Joy* perfume. I never liked my grandmother, therefore I dislike the smell of *Joy* and similar perfumes. Or, I loved my grandmother, therefore I love the scent of *Joy* and similar perfumes. Furthermore, perfumes are complex mixtures and even experts find it difficult to distinguish more than two or three components of an odor mixture. However, the specific feature(s) of a mixture that individuals recognize can vary because of their past experiences and familiarity with different odors. Therefore, one person may recognize the scent of rubber in a perfume and focus on it, but another picks out the scent of green apple in the same perfume and focuses on it. Thus, the same perfume smells different to each of them (like green apples or like rubber).

The reasons for wearing perfume have been found to vary with psychological and demographic factors. In the United States, young women are most conformist in their perfume preferences, picking fragrances that are popular or that their peers are wearing. Women in their 40s are most individualistic and choose perfumes that they personally like without much consideration for outside opinion, whereas women age 60 and above tend to choose perfumes that significant others have told them they enjoy. Mood has also been shown to influence perfume choice in that the chosen scent is felt to either have a positive effect on the wearer's mood or be consistent with the mood of the wearer at that time. Similarly, personality can influence perfume preferences depending on how one sees oneself (e.g., "dramatic" = heavy, oriental fragrances, or "sporty" = light, fresh fragrances), as can the situation or aim of the occasion (e.g., a romantic encounter or a job interview). Culture is another factor that determines the perfume qualities that are sought after. In particular, the Japanese prefer more muted fragrances and are less likely to wear perfume than are people from other industrialized countries.

With respect to physiology, it has been found that individuals with a "normal" sense of smell can vary in the number of functional olfactory receptors (OR) that are expressed in their olfactory epithelium; with a range between 350 and 400 functional receptors out of a possible 1,000. A greater number of functional ORs translates to the potential for greater sensitivity to certain chemicals. Highly intense odors are generally perceived as unpleasant. Thus, differences in OR expression may lead to differences in the predisposition to like or dislike various perfume compositions. Age influences olfactory sensitivity in the opposite direction. After age 65, about one quarter of the population has experienced a marked decrease in olfactory function. This decline in acuity continues to increase with age, which can affect both sensitivity to and the perceived quality of perfume. Receptor adaption that occurs after prolonged exposure to an odor (e.g., perfume) will cause wearers to be unable to perceive their perfume, but others around them can. Moreover, adaptation to a particular scent can cause reduced capability to perceive other fragrances through the mechanisms of cross-adaptation. All of the basic factors that affect olfactory sensation influence perfume perception in the same way.

Rachel S. Herz

See also Olfaction; Olfactory Adaptation; Olfactory Central Processing; Olfactory Receptors and Transduction

Further Readings

Baydar, A., Charles, A., Decazes, J. M., McGee, T., & Purzycki, K. (1996). Behavior of fragrances on skin. *Cosmetics & Toiletries, 111,* 49–57.

Calkin, R. R., & Jellinek, J. S. (1994). *Perfumery: Practice and principles.* New York: Wiley.

Classen, C., Howes, D., & Synnott, A. (1994). *Aroma: The cultural history of smell.* New York: Routledge.

Edwards, M. (2006). *Fragrances of the world 2006.* La Quinta, CA: Crescent House.

Herz, R. (2007). *The scent of desire: Discovering our enigmatic sense of smell.* New York: William Morrow/ HarperCollins.

Mookherjee, B. D., Patel, S. M., Trenkle, R. W., & Wilson, R. A. (1998). A novel technology to study the emission of fragrance from the skin. *Perfumer & Flavorist, 23,* 1–11.

Moran, J. (2000). *Fabulous fragrances II. A guide to prestige perfumes for women and men.* La Quinta, CA: Crescent House.

Turin, L. (2006). *The secret of scent: Adventures in perfume and the science of smell.* New York: HarperCollins.

Turin, L., & Sanchez, T. (2008). *Perfumes: The guide.* New York: Viking.

G

Gestalt Approach

Ever wonder how you can know that a melody played in one key is the same melody when it is played in another key, even though all the notes are different? This and related questions, such as how you segment a scene into figure and background, have been the subject of a branch of psychology called the Gestalt school since the beginning of the 20th century. The Gestalt approach to perception of considering how the mind organizes meaningless, elemental stimuli into meaningful global perceptions originated when Christian von Ehrenfels (1859–1932) published his pivotal paper "Über Gestaltqualitäten" (On the Qualities of Form) in 1890. This work postulated the radical view that "the whole is greater than the sum of its parts." What von Ehrenfels meant by this now-classic phrase was an extension of an earlier view postulated by philosopher Immanuel Kant (1724–1804). Kant held that the mind's representation of reality requires active processing of sensory information. Kant called our mind's representation of reality the *phenomenal world,* but reality itself the *nominal world.* That is, we perceive the phenomenal world only by filtering the outside world, which Kant called the *noumenal world,* through our minds. The noumenal world consists of "things in themselves" and can never be experienced directly, whereas the phenomenal world is created by adding intuitions and conceptions to our sensory impressions. Thus, von Ehrenfels's *Gestaltqualitäten* cannot be explained by merely associating elementary sensations, but require mental interaction with sensations.

Although there is no direct English counterpart for the German word *gestalt,* commonly employed terms are *configuration, holistic, form, structure,* and *pattern. Gestalt* describes a configuration or form that is unified. The notion that the mind actively organizes elemental stimuli from the outside world was in direct opposition to the Structuralist philosophy put forward most directly by Wilhelm Wundt (1832–1920) and championed in the United States by his student Edward Titchener (1867–1927). Structuralism postulated that all psychological facts consist of unrelated inert atoms that could only be combined by associations. They used introspection to support their claim that consciousness was composed of mechanistic associations of stimulus elements without any underlying meaning gluing them together. So, that an apple would be described as "red," "round," "hard," or "sweet," Wundt and Titchener essentially imagined psychology as a science, much like physics or chemistry, in which consciousness is a collection of identifiable parts such as the elements on a periodic table that can be separated and studied. In contrast, the Gestalt approach focused on phenomenology; that is, the study of the meaningful, intact experience not analyzed or reduced to elemental parts. A square has a unity and identity that cannot be fully appreciated by its description as four straight lines connected by right angles.

Wertheimer's Apparent Motion Phenomenon

Having taken several courses from von Ehrenfels, Max Wertheimer (1880–1943) formally founded the Gestalt approach to perception in 1912 in *Experimentelle Studien über das Sehen von Bewegungen* (Experimental Studies on the Perception of Movements), which focused on apparent rather than real motion. Wertheimer showed how two discrete lights flashed at two different locations at different times could appear to be one light that moved from the first to the second location if the spacing and timing between the stimuli were appropriately configured. He found that if the time between flashes was 200 milliseconds or longer, the observer perceived two lights flashing on and off successively, which was the case. If the interval between flashes was 30 milliseconds or less, both lights appeared to be on simultaneously. However, if the interval between flashes was about 60 milliseconds, *one* light seemed to be moving from one position to the other. This is how a theater marquee makes lights appear to move simply by flashing them on and off sequentially. This result was significant because an observer perceives neither two separate elements nor two local events, as the Structuralists would have predicted. They assumed a necessary one-to-one correspondence between perception and sensory stimulation. Instead, motion is seen where none exists. The whole pattern differs from, and supersedes analysis of, summation of the parts. Similarly, what came to be known as space-form and time-form are used by Ernst Mach (1838–1916) in *The Analysis of Sensations* (1885) to show how a melody differs from the elements that compose it. For example, transposition to a different key changes all the individual notes, yet the clear identity of the melody remains, and it remains if we change its tempo. The Gestalt school argued that we perceive our world through meaningful patterns or unified wholes.

Figure and Ground

Edgar Rubin (1886–1951) considered the two components of meaningful perceptual relations as figure and ground. *Figure* refers to the "thingness" of an object, whereas *ground* conveys the remainder of a scene, or the background. To show that figure and ground were mutually exclusive, Rubin created such stimuli as his famous reversible faces/vase (Figure 1). When the faces appear as the figure, the vase appears as the ground, but when the vase appears as the figure, the faces recede into the background immediately completely. Rubin took this result as evidence of the integration and segregation of meaningful entities that appear to exist in the world.

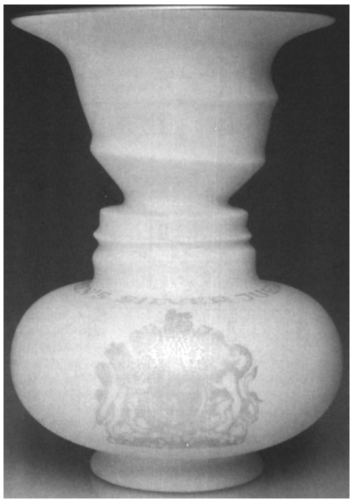


Figure 1 Figure/Ground (Reversible Faces and Vase)

Köhler and Koffka's Principles of Organization

Although over the years Gestaltists have isolated more than 100 laws that explain how perceptual information is arranged, psychologists Wolfgang Köhler (1887–1967) and Kurt Koffka (1886–1941) proposed several principles of organization that the mind evokes whenever it encounters elemental stimuli. The first and perhaps easiest to understand is the principle of proximity. Items that are close to one another in time or space lead to a perception of togetherness. Notice that in Figure 2(a), we see four columns of squares instead of four rows. This grouping occurs entirely in the mind's eye because neither configuration is more likely to be apparent given a Structuralist framework. Thus, proximity produces two effects. First, it groups or integrates elements together, which is called *belongingness,* and second, it segregates elements that do not unite in proximity—for example, across columns. This integration/segregation is thought to underlie figure-ground relations as well as group individual elements.

The second principle, *similarity,* states that similar things tend to be grouped together. Notice how, even though the columns and rows in Figure 2(b)

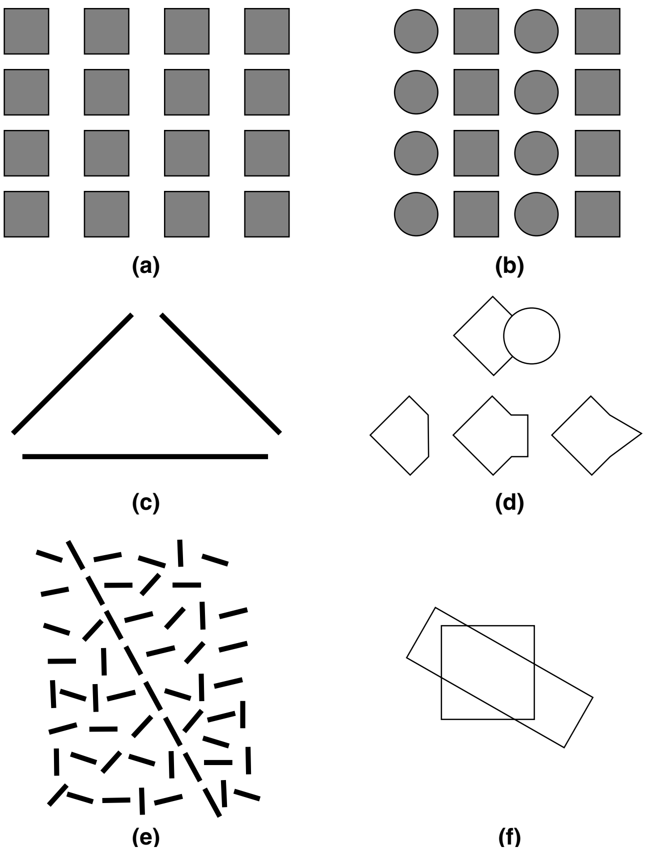


Figure 2 (a) Proximity, (b) Similarity, (c) Closure, (d) Symmetry, (e) Good Continuation, (f) Law of Prägnanz

have the same proximity, we group the stimuli into columns of similar elements: circles and squares.

A current psychophysical methodology tries to discriminate which Gestalt principle is more elementary, proximity or similarity. Increasing the spacing between the rows in Figure 2(b) can determine when the principle of proximity will prevail over the principle of similarity.

A third principle, *closure,* says that we tend to complete figures that are incomplete—we fill in the gaps. Thus, in Figure 2(c), we tend to see a triangle rather than a series of disconnected lines, entirely because of the mind's active processing.

Symmetry is a fourth principle. Figure 2(d) shows how we favor the perception of a diamond on the top image, even though part of it is occluded by a circle. Notice that none of the three possible solutions to the occluded structure (shown below it) seem likely. The ability to perceive the structure as a diamond is called *amodal completion,* meaning without the explicit use of a sensory mode, such as vision, and develops because the visual mode does not really see the occluded portion of the diamond but completes the structure amodally because of the principle of symmetry.

A fifth principle, *common fate,* states that objects that move in the same direction are perceived to be grouped together. Thus, if there is an array of dots, and half of them are moving toward the right and the other half are moving toward the left, the array will appear to be composed of two groups, a rightward and a leftward moving group.

A sixth principle is called good continuation. In Figure 2(e), we tend to group the aligned segments into a slanted line, but the other segments appear disconnected, as if randomly positioned. This introduces the most central general grouping principle of organization that Wertheimer called the Law of *Prägnanz,* which refers to the essence or ultimate meaning of an experience. According to this law, perceptual organization tends to be as good as prevailing conditions allow. Thus, we see our world in as orderly, coherent, and economical a way as conditions permit. A good gestalt is symmetrical, simple, and stable and cannot be reduced to a simpler form or made more orderly, like a soap bubble. Thus, when looking at Figure 2(f), we tend to see a square overlapping a rectangle rather than a combination of more complicated shapes. Though looking at Figure 2(f) gives us an intuitive feeling for this law, a common criticism of the Gestalt approach has been that its terms are loosely defined. For example, what does the word *good* mean in the previous definition? How do we know when a given perceptual organization is as good as possible? As the structure of stimuli varies, even slightly, a precise definition of "goodness" can become intractable.

Historical Relationship Between Gestalt Findings and Physiology

Because Gestaltists were able to produce stimuli that directly and convincingly support their principles, they argued that qualitative data be the focus of perceptual scientific inquiry instead of the more traditional quantitative analysis. This viewpoint set the Gestalt approach apart from mainstream science. During this epoch, issues such as how principles such as good continuation could be accommodated with what was believed to be known about brain physiology were considered. For example, in Figure 2(e), the observer's task might be to discriminate as quickly as possible the path of good continuation among the lines of different orientation. Each line is thought to evoke a response from a unique brain region specifically tuned to that orientation. The pattern perceived within the image is extracted because the orientation of each line that forms the larger 45-degree line has the same orientation, thus producing the largest cortical response, which allows the cortex to group like orientations into a meaningful unit.

Gestalt psychologists argued that these principles of organization reflected the physiological organization of the brain and not the processes of the mind, as Kant suggested. Köhler described this notion, called *psychophysical isomorphism,* as relating how space is structured to also have a functional order that is reflected in the distribution of underlying brain processes. Yet, he did not mean that the brain contains *pictures* of the outside world but, rather, a *functional* equivalence. Gestalt psychology once again differs from the Structuralist position that the brain is mechanistically organized to extract elements of conscious experience. Instead, Gestaltists hypothesized that a sensory stimulus enters structured fields of electrochemical forces in the brain, both modifying them and being modified by them. What we experience consciously results from this interaction. The key point is that brain activity transforms sensations and gives them emergent characteristics they otherwise would not possess. According to this historical analysis, the whole (the brain's electrochemical force fields) exists before the parts (sensations), and the whole gives the parts their meaning.

Gestalt Principles and Perceptual Research

By the 1920s, Gestalt psychology was a burgeoning enterprise through its journal, *Psychologische Forschung (Psychological Research)*. However, the Nazi rise to power in 1933 split the group before members could establish a doctoral program. They immigrated to the United States and became established at different universities, preventing a cohesive program from developing. However, the strength of their ideas and the compelling simplicity of their stimuli convinced other perceptual scientists to incorporate Gestalt theory into mainstream research. With the current rise of computer vision, many have looked for ways to train their algorithms to recognize the Gestalt grouping principles as a means to extract coherence from elemental stimulus sets, as in *top-down processing*. Thus, the Gestalt approach to perception has been reinvigorated, developing new principles, and looking to incorporate existing ones into current perceptual models.

James A. Schirillo

See also Perceptual Organization: Vision; Perceptual Segregation; Theoretical Approaches; Vision

Further Readings

Ehrenfels, C. von (1890/1988). On "Gestalt qualities." In B. Smith (Ed. & Trans.), *Foundations of Gestalt theory* (pp. 82–117). Wien, Germany: Philosophia.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace, & World.

Köhler, W. (1940). *Dynamics in psychology*. New York: Liveright.

Smith, N. K. (2003). *A commentary to Kant's critique of pure reason*. Hampshire, UK: Palgrave Macmillan.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegungen *Zeitschrift für Psychologie, 61,* 247–250.

Wertheimer, M. (1938). Gestalt theory. In W. D. Ellis (Ed. & Trans.), *A source book of Gestalt psychology* (pp. 1–11). London: Routledge & Kegan Paul. (Original work published 1925)

GRASPING

See Reaching and Grasping

GUIDANCE SYSTEMS FOR BLIND PEOPLE

Human wayfinding involves two different functions: (1) maintaining orientation within the large-scale environment for the purpose of navigating to remote destinations, and (2) sensing of the immediate environment, including paths and obstacles, for the purpose of moving along desired routes. In connection with the training of wayfinding skills in blind people, these functions are referred to as *orientation* and *mobility*. Guidance systems have been developed to assist blind people with orientation and mobility. This entry describes a number of these guidance systems.

Visually impaired people are at a disadvantage when it comes to wayfinding. For mobility, sensing of the near environment for obstacles and paths is effortful and error prone, even with a long cane and auditory cues such as sound reflections from surfaces. For orientation, especially within unfamiliar environments, blind people lack much of the information needed for planning routes to destinations, for taking detours while en route, and for keeping track of their movement relative to the destination.

Early research and development of electronic travel aids (ETAs) focused on mobility, the ultimate goal being to supplement or even replace the long cane and the guide dog. The primary concern was obstacle avoidance. Electronic obstacle avoiders used today include devices that detect nearby obstacles using laser and ultrasonic sensing and then inform the user of their locations by way of auditory or vibratory displays. Some useful devices have been developed, but adoption of these devices by the blind population has been limited.

More recently, research and development of ETAs has shifted from mobility to orientation. There are basically two approaches: add electronic location identifiers to the environment or provide the traveler with an electronic device that locates the traveler within the environment. An example of the first approach is the Talking Signs© system of remote signage. Talking Signs are infrared transmitters positioned throughout the environment, such as in shopping centers and public transportation terminals. Infrared signals from the transmitters convey speech information about the identity of the site at which the transmitter is located (e.g., a bus stop). A blind person traveling through an environment picks up signals from the Talking Signs when aiming a small handheld receiver toward any one of them and, consequently, hears speech indicating the identity of the site. An obvious drawback of this approach is the cost of deploying and maintaining so many transmitters.

The second general approach for providing orientation information is best exemplified by guidance systems that use the global positioning system (GPS). A GPS receiver uses signals from satellites to compute its position with an accuracy suitable for pedestrian travel. A limitation of GPS is that positioning is poor indoors and somewhat degraded by tall buildings and dense foliage out of doors. In these situations, GPS guidance systems need to be supplemented by some other means of position tracking, such as Talking Signs or updating the traveler's position using velocity or acceleration sensors that measure the traveler's motion.

A GPS guidance system consists of four components: the GPS receiver for providing current location, the spatial database of streets and environmental sites, the system software for path planning, and the user interface, which displays guidance information using auditory or tactual information and allows user input to control the system. There are now several commercially available GPS guidance systems for blind people, notably Braille Note GPS©, Mobile Geo©, StreetTalk©, and Trekker©. These come equipped with detailed databases of street networks and points of interest (restaurants, museums, businesses, etc.). The effectiveness of these guidance systems in allowing blind people to travel independently through familiar and unfamiliar environments makes these systems one of the true success stories in the field of sensory substitution.

As part of the user interface, these commercial products all display guidance information to the traveler using verbal information that is communicated by way of synthesized speech or an electronic Braille display. In the former case, the guidance system speaks the traveler's current position using a street address or street intersections and provides verbal directions for the route to be traveled. A possible improvement over speech display is a spatial display that uses perceptual information to directly indicate environmental locations, such as the destination, the next turn point along the route, and points of interest.

A group at the University of California, Santa Barbara (UCSB), has done extensive research on spatial displays as part of the user interface for guidance systems. With spatial displays, the traveler experiences environmental locations from a first-person perspective. Foremost is a display that uses virtual sound, as conveyed by a virtual acoustic display. The idea is that as the blind person travels through the environment, he or she hears the names of buildings, street intersections, and so on, spoken by a speech synthesizer and appearing to come from loudspeakers at the appropriate locations. To create such an impression, the virtual acoustic display takes a monaural audio signal, such as synthesized speech, and converts it into signals delivered by earphones, signals that contain the direction and distance cues of normal listening.

The direction to a real sound source is indicated by two binaural cues, interaural time difference (ITD) and interaural level difference (ILD). ITD is non-zero whenever one ear is closer to the source than the other ear, resulting in different travel times of the sound to the two ears. ILD is non-zero whenever one of the ears is facing away from the source and is partially blocked by the intervening head, resulting in different sound intensities at the two ears. Because the computer modifies the signals to the two ears according to information coming from an electronic compass mounted on the user's head, the sound from an environmental location appears to remain fixed in direction relative to the person's body, despite head rotations. The distance to a real source is indicated by a number of cues, primarily overall sound intensity and the ratio of the intensity of the sound that arrives first (traveling directly from the source) to that of intensity of the reflected sound that arrives shortly thereafter. The virtual acoustic display simulates these and the secondary distance cues. Together, the simulated distance and direction cues create a moderately realistic impression of external sound sources.

Another type of spatial display makes use of a handheld pointing device similar to the handheld receiver used with Talking Signs. When the traveler scans with the device, which contains an electronic compass, and locates an environmental location within the spatial database by pointing in its physical direction, the computer presents synthesized speech about the identity of the location and its distance. Whereas a real Talking Signs receiver actually picks up information from an infrared transmitter, this type of spatial display only mimics the use of such a receiver.

Research on route guidance by the UCSB group has shown that both the virtual acoustic display and the handheld pointer are well liked by blind travelers and lead to efficient traversal of routes, sometimes approaching the efficiency of sighted travelers. This research has also shown that synthetic speech conveyed by virtual sound, when compared with conventional synthetic speech, is preferred by blind travelers, results in faster route traversal, and is less cognitively demanding. Given the effectiveness of spatial displays, future guidance systems will likely offer either virtual sound displays or the handheld pointer as alternatives to speech display.

Jack M. Loomis

See also Auditory Localization: Psychophysics: Echolocation; Low Vision; Sensory Restoration and Substitution; Visual Disorders: Blindness

Further Readings

Giudice, N. A., & Legge, G. E. (in press). Blind navigation and the role of technology. In A. Helal, M. Mokhtari, & B. Abdulrazak (Eds.), *Smart technology for aging, disability, and independence: Computer and engineering for design and applications*. New York: Wiley.

Golledge, R. G., Klatzky, R. L., Loomis, J. M., Speigle, J., & Tietz, J. (1998). A geographic information system for a GPS based personal guidance system. *International Journal of Geographical Information Science, 12,* 727–749.

Hatakeyama, T., Hagiwara, F., Koike, H., Ito, K., Ohkubo, H., Bond, C. W., & Kasuga, M. (2004). Remote infrared audible signage system. *International Journal of Human-Computer Interaction, 17,* 61–70.

Loomis, J. M., Golledge, R. G., Klatzky, R. L., & Marston, J. R. (2007). Assisting wayfinding in visually impaired travelers. In G. Allen (Ed.), *Applied spatial cognition: From research to cognitive technology* (pp. 179–202). Mahwah, NJ: Lawrence Erlbaum.

Welsh, R., & Blasch, R. (1980). *Foundations of orientation and mobility.* New York: American Foundation for the Blind.

H

HALLUCINATIONS AND ALTERED PERCEPTIONS

Sensory percepts may sometimes occur in the absence of adequate sensory stimulation. The most prominent examples of nonveridical percepts are hallucinations. They can be pathological, but can also affect healthy individuals. Nearly 10% of the general population reported having experienced an unexplained perception; about 3% reported having heard a voice. Hallucinations may result in secondary *delusions*—inaccurate explanations of what is happening. This entry discusses hallucination in healthy people, in mental disorders, and in other medical conditions; other altered perceptions; nonveridical percepts in everyday life; and neuronal mechanism of nonveridical perceptions.

Hallucinations in Healthy People

Healthy people can experience hallucinations. For instance, the hearing of a family member's voice is not uncommon among recently bereaved people. These hallucinations become less frequent and cease over weeks or months. They are comforting and benign.

Some people take hallucinogenic drugs, such as LSD or mescaline, with the clear intention of inducing hallucinogenic experiences. Hallucinations may also occur with some medically prescribed drugs, for example, the anesthetic drug ketamine, or the drug used in the treatment of Parkinson's disease, levodopa. In drug withdrawal states, particularly withdrawal from alcohol, hallucinations are common.

Visual hallucinations have been reported during sensory deprivation in healthy individuals, either as a result of solitary imprisonment ("prisoner's cinema"), or during prolonged blindfolding. Hallucinations can occur in pilots during long night flights. Hallucinations also occur during snowstorms, suggesting that absence of patterned stimulation is more likely to produce visual hallucinations than is light deprivation.

Hallucinations in Mental Disorders

Auditory hallucinations usually mean hearing voices. However, nonverbal auditory hallucinations do occur and include clicking and mechanical noises, muttering or mumbling, and music. In musical hallucinations, the patient often hears a complete piece of music. Auditory hallucinations are most common in psychotic disorders such as schizophrenia, but can occur in mood disorders, organic mental disorders, and drug-induced states. Auditory hallucinations occasionally occur in the elevated phase of bipolar disorder and in severe (psychotic) depression.

Visual hallucinations can be classified into simple and complex hallucinations. Simple hallucinations mostly consist of dots, lines, geometric shapes, and moving patterns. Complex hallucinations include the occurrence of other people, animals, and more rarely objects such as cars or tables. Visual hallucinations may occur with such brain disorders as tumors, multiple sclerosis, and dementia. They occur more frequently than auditory hallucinations in the organic mental disorders. In some types of epilepsy, visual hallucinations may form complex scenes. In schizophrenia, visual hallucinations are often indistinct or distinct figures, often humanoid, standing to one side of the patient.

Tactile hallucinations include the experience of being touched, or of a crawling sensation under the skin. These are common in drug withdrawal states, but may occur in schizophrenia. *Somatic hallucinations* are the sensation of things happening inside the body, such as organs moving from one part of the body to another. These are rare, but may occur in schizophrenia. *Gustatory hallucinations,* the hallucinations of taste and smell, are common in such medical conditions as epilepsy, but can rarely occur in schizophrenia.

Hallucinations in Other Medical Conditions

In migraine, a common condition involving people without a mental disorder, visual hallucinations are usually experienced just before a migrainous attack ("aura"). Sometimes the aura can occur in the absence of headache. The percepts during the migraine aura take the form of scintillating, jagged lines, called *migraine fortification spectra,* which involve one half of the visual field and expand from the central field toward the periphery. The specific evolution in time of the percepts occurring during the migraine aura suggests that they might be related to biochemical changes that invade portions of the topographically organized visual cortical areas ("spreading depression"). Migraineurs are believed to show deficient cortical inhibitory processes.

In Charles Bonnet syndrome, which affects mostly healthy elderly, visual hallucinations are associated with blindness and social isolation. Common causes are degenerative ocular diseases such as macular degeneration, cataract, or glaucoma.

Other Altered Perceptions

Illusions are perceptions associated with an outside stimulus, but the stimulus is wrongly interpreted. For example, lapping water may be heard as laughter. One condition that most commonly causes illusions is delirium tremens, the disturbed state occurring with alcohol withdrawal. *Heightened perceptions* include sounds that seem unnaturally clear, loud, or intense; colors that appear more brilliant or beautiful; or details of the environment that seem to stand out in a particularly interesting way. *Changed perceptions* include changes that are perceived in the shape or size of people and inanimate objects in the environment. Changes may continue while the patient watches. Heightened and changed perceptions may occur in several psychotic disorders, but also in the nonpsychotic anxiety disorders. *Nonveridical visual percepts,* consisting of idiosyncratic, moving, and colored distortions of the viewed images, are experienced in the chronically suppressed eye of subjects with neurodevelopmental disorders such as strabismic amblyopia. These distorted visual percepts might be related to the prolonged disuse of the cortical structures connected to the amblyopic eye.

Nonveridical Percepts in Everyday Life

Flashes of light, called *phosphenes,* can be induced by mechanical, electrical or magnetic stimulation of the retina or the brain. They are well localized in the visual field and might appear in different colors. Phosphenes are usually amorphous, but can develop into geometric patterns; their appearance depends on the stimulated location.

Visual phantoms were described for the normally occurring phenomenon of filling-in. Illusory perception of apparently moving stimuli is perceived along the illusory movement trace, in the absence of an actual stimulus.

In *mental imagery,* subjects willfully summon particular images or scenes. Imagery and perception were suggested to share similar resources. Indeed, activations in the primary visual cortex have been reported during visual imagery.

Neuronal Mechanisms of Nonveridical Perceptions

One feature common to all neurobiological explanations of hallucinations is the presence of increased neural activity. Based on the observation that patients with eye disease experience the same pathologies of visual perception as did patients with cerebral lesions and, under certain circumstances, normal subjects, researchers suggested that the pathological percepts related to hallucinations can be interpreted as "positive" symptoms of a disturbed brain function.

Ruxandra Sireteanu

See also Amblyopia; Auditory Imagery; Migraine; Nonveridical Perception; Olfactory Imagery; Sleep and Dreams; Visual Filling In and Completion; Visual Imagery

Further Readings

Burke, W. (2002). The neural basis of Charles Bonnet-hallucinations: A hypothesis. *Journal of Neurology, Neurosurgery & Psychiatry, 73,* 535–541.

Farah, M. J. (1988). Is visual imagery really visual? Overlooked evidence from neurophysiology. *Psychological Review, 95,* 307–317.

Ffytche, D. H., & Howard, R. J. (1999). The perceptual consequences of visual loss: "Positive" pathologies of vision. *Brain, 122,* 1247–1260.

Manford, M., & Andermann, F. (1998). Complex visual hallucinations: Clinical and neurobiological insights. *Brain, 121,* 1819–1840.

Ohayon, M. M. (2000). Prevalence of hallucinations and their pathological associations in the general population. *Psychiatry Research, 97,* 153–164.

Schultz, G., & Melzack, R. (1991). The Charles Bonnet syndrome: "Phantom visual images." *Perception, 20,* 809–825.

HAPTICS

The word *haptic* refers to the wealth of perceptual experience obtained from skin, muscles, tendons, and joints, especially through manual exploration. Haptic perception encompasses multiple distinct input sensations, resulting from different types of neural structures that receive stimulation from the world and convey it to the brain. Inputs from sensors within the skin are called *cutaneous,* or sometimes *tactile,* whereas those from sensors in muscles, tendons, and joints are called *kinesthetic.* This entry discusses several aspects of haptic perception.

Neural Basis of Haptic Perception

The foundation of haptic perception lies in sensory receptors, populations of neurons that convert events impinging on the body to electrical signals sent to the brain. A receptor consists of the axon of a nerve fiber, which in some cases originates in a specialized ending. Each type of haptic receptor (or sensor) responds to a particular kind of information and, hence, conveys a particular kind of interaction with the world.

The cutaneous subsystem of touch begins with receptors that lie within the skin. Some are found in the outer layer, the *epidermis,* whereas others lie more deeply in an underlying layer, the *dermis.* Some of these receptors are responsive to mechanical interactions with the skin, that is, to forces that press or vibrate against it. These *mechanoreceptors* can be categorized by the area of skin or *receptive field* that makes them active and by the duration of their activity once they are activated. Receptors that are spatially selective, that is, that selectively respond to stimulation only within a small region of skin, tend to be located near the surface in the epidermis, whereas receptors that respond to stimulation across a less-precisely bounded area of skin lie deeper, in the dermis. Receptors that continue to fire during sustained pressure of the skin are called slow-adapting, whereas those that cease until stimulation begins again are called fast-adapting.

These different receptor response characteristics inform us that the skin contains different signaling mechanisms depending on the type of stimulus. Light fluttering is sufficient to activate fast-adapting fibers near the skin surface, whereas vibrations that arise when the skin contacts an object activate deeper fast-adapting receptors. Slow-adapting fibers near the skin surface, by virtue of their small receptive fields, provide a spatial map of pressure on the skin, allowing a blind person to read a Braille character or an embossed sign.

Other cutaneous receptors, called *thermoreceptors,* respond to warmth or cooling of the skin. Unlike mechanoreceptors, these neural fibers lack specialized endings. Warmth fibers fire when the skin temperature rises as, for example, when you go outside on a sunny day. Cold fibers become active when your skin temperature decreases, as when you hold an ice cube. Warmth and cold fibers are useful for identifying the material from which an object is made. In a typical room, objects that do a good job of conducting heat away from your skin (e.g., steel) tend to feel cooler than do those that conduct heat poorly (e.g., wood).

Haptic receptors responding to mechanical stimulation are found in muscles, tendons, and joints, as well as in skin. Collectively, these constitute the kinesthetic system, telling you where your limbs are in space and their movements, along with the force your muscles are exerting. Kinesthetic receptors are essential for monitoring the activities of your body, whether you are playing tennis, walking upstairs, or simply standing.

Although it is an unpleasant subject, pain is also conveyed by fibers in the skin, muscles, tendons, and joints. These fibers, called *nociceptors,* are aroused by extremes of normal stimulation, whether pressure (a sharp pinch), heat, or cold.

Recent research has indicated the existence of one additional type of fiber that plays a role opposite to the nociceptor by conveying pleasant or emotional touch and is most typically found in hairy skin. An optimal stimulus for this type of receptor appears to be light stroking or petting.

Mechanoreceptors, thermoreceptors, and nociceptors are at the beginning of a neural chain of events culminating in the activation of neurons in the brain. There are two major pathways from the peripheral receptors to the brain, corresponding to a division in evolutionary history. Roughly speaking, pain and temperature signals are sent along an older, slower pathway via the spinal cord, whereas sensations arising from mechanical interactions are sent more quickly via a newer system that enables the rapid initiation of action. Both neural pathways ultimately arrive at an area within the brain called the primary somatosensory cortex. If you place your right hand on the top of your head, with your middle finger crossing the center line, your index finger will lie approximately over the somatosensory area of the brain, and your fourth finger will lie over the area that conveys commands that go down to the muscles. This proximity of sensory to motor probably reflects the close functional relationship between haptic perception and action.

Haptic Perception in Relation to Action

The connection between haptic perception and action is a two-way street. First, haptic perception supports action. Anyone who has ever tried to manipulate an object with very cold hands knows that touch is integral to fine control of grasping and manipulating, for example, inserting a key into the lock in a door or simply holding the key without dropping it. In the laboratory, anesthetizing the skin on the hands has been used as an effective method for showing that tactile sensing is essential for action. People who lack skin sensations cannot maintain an object in a grasp. Measurements of neural signals from mechanoreceptors have shown just why skin sensation is so important for grasping: Fast-adapting fibers near the skin surface become active when an object just begins to slip from the fingers, and their signals are conveyed so rapidly that the grasp can be corrected before the object actually drops.

On the other side of our two-way street, action provides support for haptic perception. Want to find out what an object is made of? Act on it. If, for example, you want to feel whether a surface is rough or smooth, you are more likely to rub your fingers across it than to statically contact it. It has been shown that as you rub, the sideways movement of your fingertip across the surface enhances the activity in mechanoreceptors that convey information to the brain about how rough it feels.

More generally, specific links have been found between the way you manually explore (type of hand action) and the specific object property that you feel. For example, to find out if an object is soft, you press or bend it; to find out if an object is heavy, you pick it up and heft it. Each of these actions, termed *haptic exploratory procedures,* is a characteristic movement used to evaluate the associated property and has proved to be optimal for evaluating that property. Together, haptic perception and action form a powerful team, allowing us to learn about objects around us and to use this information for many different purposes.

Haptic Perception of Objects and Their Properties

Learning about the properties of objects by touch further enables us to identify them without looking. Although perhaps surprising, it is easy to demonstrate that most common objects placed in the hands can be identified within a couple of seconds. While en route to identification, people are often observed to rub, press, finger, and in general, handle the objects to execute informative exploratory procedures. Merely grasping and lifting an object is often sufficient to name the object with its most common name, such as "pen." More extensive exploration is used to identify the type of object with still greater precision, as in a "ballpoint pen."

Studies of how the haptic system achieves rapid object identification reveal that haptic processing is highly complementary to vision. Vision conveys information about the geometric properties of objects—their curvature, elongation, or size. Touch is far less fast or precise at determining an object's geometry, but it excels at ascertaining its material properties—hardness, roughness, stickiness, perceived warmth or coolness, and weight, for example.

The division of labor between haptic and visual perception is apparent from a number of scientific findings. One comes from observing whether people touch objects if vision is also available for identifying and discriminating among them. Faced with a difficult question about geometric properties, such as comparing the size of a coin and a circular battery, people will look closely but rarely touch. But when a difficult question about material properties is posed, for example, comparing the roughness of an egg and a sheet of rag paper, people tend to touch the objects, typically with the appropriate exploratory procedure.

When people both look at and touch objects, the two senses tell them about different surfaces. You can't see through your fingers as you touch an object. Pick up your cell phone and look at it, and you are likely to find yourself looking at the keypad but feeling the back. When participants in a controlled experiment are asked to haptically identify objects they have previously seen, they do best when the test objects are reversed so that participants can touch the object surfaces they previously saw. Conversely, if tested visually with objects they previously felt, they had better be shown the back side.

What about "objects" such as Braille symbols, which vary meaningfully in their geometric pattern but must be recognized by touch? The Braille alphabet has been constructed so that a symbol fits within the fingertip and, hence, can be felt in its entirety at one time. The scaling of Braille to the finger size bypasses one of the principal problems with determining the shape of an object by touch, namely, that it must be explored piecemeal over time, and then somehow mentally pieced together. Despite the advantage of scaling, however, reading Braille remains a difficult skill to master. Even with extensive exposure, sighted teachers of Braille read the symbols with their eyes, rather than with their hands.

The human face is another type of object that is amenable to haptic recognition. People show considerable success at identifying specific live faces by touch, and at haptically recognizing live facial expressions of emotion. We should not be surprised at these findings, however, when we consider that facial features and their alterations under different emotional expressions are haptically, as well as visually, informative. Joe has high cheekbones and a broad nose; Jane's eyes crinkle at the corner as she smiles. The structural and textural properties that enable us to visually recognize Joe or Jane, or to see them smile or frown, also leap to life under the hands.

Haptic Perception of Space

Receptors in muscles, tendons, and joints provide information about how our limbs are disposed in space, enabling us to reliably repeat movements, such as finding the gearshift in our car. (We become aware that we have this so-called muscle memory when we rent or borrow someone else's car and reach in the wrong place.) Kinesthetic receptors do not, however, readily enable us to convert our limb positions into a "mental map" of the places we have previously touched. Close your eyes, reach out, and feel the objects around you. It is likely you can then make a sketch of their spatial arrangement. But just how accurate is it? The perception of spatial layout by touch turns out to be error-prone, and in ways that are informative.

Here is a nice demonstration. Without looking, put your left and right index fingers down on a tabletop separated by some arbitrary distance, and ask yourself how far apart they are. You are likely to make a sizable error, on the order of 10% or more. That error in estimating the distance between the two fingers will increase dramatically if you first put them together, then separate them by having one move along some wiggly path to its new stopping place, before making the response. The longer the path used to separate the fingers, the greater the judged straight-line distance between them tends to be. Such results suggest that when using haptic exploration, the time taken by the moving hand is used as one way to estimate the distance between two points in space.

Errors in haptic space perception can be demonstrated by another task, adjusting the orientation of two bars so that they feel parallel. Place a pencil off to your left side on the table top at an arbitrary angle, put your hand on it, and then try to rotate a pencil off to your right side so that the two feel parallel. Depending on the positions of the pencils and the angle that is to be matched, such judgments can exhibit substantial misalignments. In some circumstances, people have been observed to place the two pencils so that they form an angle that is close to 90 degrees, rather than parallel!

Haptic space perception demonstrates a number of phenomena in common with visual perception, including illusions. For example, the horizontal/vertical illusion—the tendency to judge the vertical bar of a ⊥ shape as longer than the horizontal—is found when that shape is touched, as well as when it is seen. Matching the orientation of a touched bar is less accurate when it lies along an oblique (slant) than when it is horizontal or vertical. The active nature of haptic perception brings to these phenomena some influences not found in vision: The strength of the illusion may depend on how the object is explored, for example, whether the arm movements are toward and away from the body compared with arcs that encircle the observer's body.

Applications of Haptic Perception

A haptic interface consists of a hardware system that delivers forces and vibrations to the fingers or hands of an operator, directed by software that programs the virtual characteristics (i.e., object shapes, sizes, textures, and compliances) at any moment in time. Typically, the user interacts with the device by holding a handle, or inserting a fingertip into a thimble-like holder, and explores the virtual world much as he or she would explore the world of physical objects and surfaces with a probe held in the hand. As the user moves through the workspace of the device (or is guided passively by means of a predesignated program), a force or vibration is computed according to interface software and delivered to the user. Contact with a virtual wall, for example, corresponds to a strong resistive force that the user encounters when moving the manual guide to its preprogrammed location.

The utility of a haptic interface is limited by many engineering and computer factors, including the device's spatial resolution, update speed, mechanical structure, and algorithms for rendering and control. To effectively design haptic interfaces for virtual environments and remote operation, it is critical to match the hardware and software features to the capabilities and limitations of the human operator. The scientific study of human haptics is an integral part of the design and evaluation of an interface's function. Conversely, interfaces have enabled new types of basic research on haptic perception by allowing the creation of surfaces and objects that would be unfeasible for physical manufacture, if not physically impossible. Virtual surfaces and virtual objects of arbitrary shape, size, and compliance can be rendered. Flat surfaces can be made to simulate resistive forces as if they were curved. Basic research on haptic perception is burgeoning with the help of these new facilities for creating new experimental tools.

As technology improves and new devices, both specialized and general-purpose, become available, the range of applications promises to be highly diverse. Already commercial and prototype interfaces have been developed to provide novice training in medical procedures (e.g., lumbar puncture, surgical repair), to create virtual environments for medical rehabilitation, and to control the directional movements of vehicles designed for planetary exploration.

Roberta L. Klatzky and Susan J. Lederman

See also Action and Vision; Cutaneous Perception; Multimodal Interactions: Visual–Haptic; Perceptual Development: Touch and Pain; Reaching and Grasping; Virtual Reality: Touch/Haptics

Further Readings

Burdea, G. C., Lin, M. C., & Tachie, S. (Eds.). (2005, November/December). Special issue on haptics, virtual and augmented reality. *IEEE Transactions on Visualization and Computer Graphics, 11*(6).

Heller, M. A., & Ballesteros, S. (2006). *Touch and blindness: Psychology and neuroscience.* Mahwah, NJ: Lawrence Erlbaum.

Johnson, K. O. (2002). Neural basis of haptic perception. In H. Pashler & S. Yantis (Eds.), *Stevens' handbook* *of experimental psychology: Vol. 1. Sensation and perception* (3rd ed., pp 537–583). New York: Wiley.

Jones, L., & Lederman, S. J. (2006). *Human hand function.* Oxford, UK: Oxford University Press.

Katz, D. (1989). *Der aufbau der tastwelt (The world of touch)* (L. Krueger, Trans.). Mahwah, NJ: Lawrence Erlbaum. (Original translated 1925)

Klatzky, R. L., & Lederman, S. J. (2007). Object recognition by touch. In J. Rieser, D. Ashmead, F. Ebner, & A. Corn (Eds.), *Blindness and brain plasticity in navigation and object perception* (pp. 185–207). Mahwah, NJ: Lawrence Erlbaum.

Lederman, S. J., & Klatzky, R. L. (Eds.). (2007). New directions in touch [Special issue]. *Canadian Journal of Experimental Psychology, 61(2),* 169–275.

Lin, M., & Otaduy, M. (Eds.). (2008). *Haptic rendering: Foundations, algorithms, and applications.* Wellesley, MA: A K Peters.

Newell, F. N., Ernst, M. O., Tjan, B. S., & Bülthoff, H. (2001). Viewpoint dependence in visual and haptic object recognition. *Psychological Science, 12*(1), 37–42.

HEARING AIDS

Approximately 30 million U.S. citizens have a magnitude of hearing loss that is sufficient to benefit from hearing aids. Unfortunately, for a variety of reasons, only 5 to 6 million U.S. citizens currently wear hearing aids. This entry reviews the basic function of hearing aids and provides information on some of the latest advances in hearing aid technology.

Basic Functions

Hearing aids consist of a *microphone(s)* that converts the incoming acoustic signal to an electrical signal that is a duplication of the acoustic spectrum of the incoming signal. This conversion is followed by an *amplifier* that magnifies the signal received from the microphone. This stage is typically followed by a *multichannel digital signal processor* (DSP) that manipulates the amplified signal to match the magnitude and configuration of the hearing loss of the aided ear(s). Finally, the digitally processed amplified signal is forwarded to a *receiver* that reconverts the amplified electric digitally processed signal back to an acoustical signal and sends it to the ear in a variety of ways.

Technological Advances

In the past decade, significant advances have been introduced to hearing aids, and these advances have accelerated so much that manufacturers introduce major changes every 3 to 6 months whereas in the past major changes occurred every couple of years. Current advances include *multichannel signal processing.* That is, the processing chip within the hearing aid divides the incoming acoustic signal into as few as 2 channels (low and high frequency) or as many as 20 channels (independent narrow frequency ranges). When divided in this manner, the incoming signal can be manipulated independently in each channel for improved control of amplification, noise reduction, and the management of feedback. Examples of the use of multichannel signal processing include the following: (a) digital signal processing, (b) feedback management to eliminate or dramatically reduce the presence of feedback and allow for greater available amplification, (c) noise reduction for improved listening comfort in noisy environments, and (d) automatic adaptive directional microphones for improved recognition of speech in noisy environments.

This type of microphone design contains two microphones where one is forward facing and the second faces the rear. In this design, the signal processor within the hearing aid will automatically activate the forward facing omni-directional microphone when there is a single talker in a quiet listening situation. Then the signal processor within the hearing aid will automatically activate the second rear-facing microphone when the processor detects a second signal arriving from the side or back of the listener and when the level of the rear signal is greater than the level of the signal arriving from the front. The action of the rear facing microphone is to attenuate the sound from the side or rear so that the signal arriving from the front is amplified to improve the signal to noise ratio. Advances in directional microphones have greatly improved user satisfaction and benefit with hearing aids in noisy listening environments, and this technology is improving rapidly.

Other advances include *expansion* to reduce the amplification of annoying low-frequency energy (circuit noise of the microphone and low level environmental noise) for patients with normal hearing in the low frequencies; *data logging* that allows audiologists to read from the fitting software how the hearing aid is being used and store information about the typical listening environments of the patient; *multiple memories* that allow the audiologist to place into the hearing aid different frequency/gain responses for specific listening situations such as watching television or communicating on the telephone; *wireless Bluetooth technology* for connecting the hearing aids to other communication devices (cell phone, FM listening devices, MP3 devices, etc.) or allow for "cross communication" between the right and left hearing aids so a change (i.e., manipulating the volume control, changing programs, etc.) in one hearing aid automatically occurs in the other hearing aid; and *slim tube open fits* that are small behind-the-ear hearing aids using a narrow plastic tube or wire and inserted into an open earmold to send the amplified sound to the ear canal. The use of an open earmold eliminates the occlusion effect (sensation of the head in a barrel) and provides excellent high frequency amplification.

Additional advances include *middle ear implantable hearing aids* for patients with severe hearing loss in the inner ear and a *bone anchored hearing aid* implanted directly into the mastoid for patients with hearing loss related to the middle ear or for patients with unaidable hearing in one ear and normal hearing in the opposite ear.

Another method of amplification that is independent of hearing aids and is highly effective in improving the signal-to-noise ratio is *hearing assistive technology* (HAT). This type of technology can be combined with hearing aids or can be used alone. Examples of HAT include personal FM systems that can be used alone for one-on-one communication or coupled to hearing aids using direct auditory input or via a telephone coil (telecoil) transducer available in most hearing aids. When used in isolation or in combination with hearing aids, FM systems can provide a hearing impaired listener with a level of performance in noise that is better than a listener with normal hearing. Other examples of HAT include amplified telephones, vibrating alarm clocks, alerting systems, cell phone devices, and infrared listening systems.

Modern hearing aids can be programmed for patients having normal hearing in the low frequencies with hearing loss restricted to the higher frequencies. This type of hearing loss is most applicable to slim tube open ear technology. Hearing aids can also be programmed for patients having severe to profound hearing loss at the low and high frequencies. Almost all patients can receive significant benefit and satisfaction from modern hearing aids, but one of the determinants to the magnitude of user satisfaction and benefit depends on the method used to fit the hearing aids to the patient. One method is to enter the results from the hearing test into a manufacturer's software and download the resulting manufacturer "first fit" into the hearing aids without external verification. Another method, which many audiologists consider better, is to verify, using objective probe microphone measures, that the aided frequency response meets a valid prescriptive target based upon the results from the hearing test. Probe microphone measures require a piece of equipment that places a thin plastic tube, coupled to a small microphone, in the ear canal about 4 to 6 millimeters (mm) from the eardrum. The output from the microphone is fed to a computer that analyzes the measured frequency response from the hearing aid placed in the ear canal. This measured frequency response is compared with a validated prescribed frequency response that is generated from the patient's hearing loss entered into the computer of the equipment. If the measured frequency response does not "meet" the prescribed frequency response, the audiologist will enter the software that programs the hearing aid and change the programmable parameters until the measured frequency response arrives close to the prescribed frequency response. Probe microphone measures are used to verify that the measured frequency response for a soft input level (about 50 decibels sound pressure level [dB SPL]) is audible, an average input level (about 65 dB SPL) is comfortably loud, and a loud input level (about 80 dB SPL) is not uncomfortably loud. For each of these three input levels, a different frequency response is generated. The goal is for the measured frequency response for the three input levels to match the prescribed frequency response for the three input levels. For maximum patient benefit and satisfaction, all hearing aid fittings should be verified using objective measures such as probe microphone.

Future Directions

As mentioned earlier, hearing aid technology is advancing at a rapid rate. Advances in the near future will allow for immediate communication between hearing aids and all Bluetooth enabled devices at the touch of a button. Also, hearing aids will cross communicate and make decisions about differences in time, intensity, and phase between the two ears as the input signal(s) arrive off center (i.e., to the left or right of the patient). This will allow for improved performance in noise and better localization. Finally, microphones and receivers will continue to reduce in size, and hearing aids will become increasingly cosmetically appealing.

Michael Valente

See also Ageing and Hearing; Audiology; Audition; Audition: Disorders; Auditory Imagery; Auditory Processing: Peripheral; Auditory Scene Analysis; Perceptual Development: Hearing; Perceptual Development: Speech Perception; Sensory Rehabilitation; Sound Reproduction and Perception; Sound Stimulus; Speech Perception; Timbre Perception; Virtual Reality: Auditory; Word Recognition

Further Readings

Dillon, H. (2001). *Hearing aids.* New York: Thieme Medical.

Valente, M. (Ed.). (2002). *Hearing aids: Standards, options and limitations.* New York: Thieme Medical.

Valente, M. (Ed.). (2002). *Strategies for selecting and verifying hearing aid fittings.* New York: Thieme Medical.

HEARING IMPAIRMENT

See Audition: Disorders

HUMAN–MACHINE INTERFACE

Few aspects of modern life have gone unaffected by machines and computers, devices that amplify our ability to perform useful tasks. Whether mechanical or mental aids, these devices are only as useful as our ability to operate them, which can often be directly attributed to the efficacy of the human–machine interface, the means by which users and the system communicate with each other. The human–machine interface includes the hardware and software used to translate user input into commands and to present results to the user. The human–machine interface is also referred to as the human–computer interface or simply the user interface. This entry describes usability, input and output, and evolution of the human–machine interface.

Usability

Usability of the human–machine interface is the degree to which the design makes using the system effective, efficient, and satisfying. Because it is generally easier to engineer and influence the way the machine works than trying to change the human, the driving philosophy has been to build interfaces based on deep understanding and appreciation of human physical, mental, and behavioral capabilities, rather than assuming the user will learn and adapt to the system, as was the case in the early days of computing.

In the classic human–machine model, the human and the machine are treated as information-processing devices. Just like humans, computers are assumed to sense information encoded as inputs, compare, choose, and formulate appropriate responses, and then communicate them as outputs. In this model, the outputs from one component of the system feed into the inputs of the other. For example, the output from the human, such as moving a mouse to communicate intentions, forms the input to the machine. Because we have traditionally interacted with the external world through our physical bodies, most computer input mechanisms require performing some form of motor activity, be it moving a mouse, hitting buttons, or speaking.

Input and Output

The fields of ergonomics and human factors engineering have focused large efforts on designing input devices and techniques to accommodate limits of the human user and even to exploit these limits. Common computer input devices include pointing devices such as the mouse, trackball, joystick, and specialized three-dimensional trackers as well as various keyboards and keypads. Other "natural" modalities are starting to embed traditionally human perceptual and cognitive abilities into computers before they can sense and interpret richer input. For example, speech input, vision, or touch-based gesture recognition and handwriting recognition are starting to become quite popular. As a further example of this, some computers can now implicitly infer such things as emotion or cognitive load from video streams of a user's face or from physiological signals such as heart rate. Given the wide assortment of input available, the actual choice of device and application can often be based on the task, users, or environments in which they are to be used.

Driven largely by the needs of people with physical disabilities, researchers have also begun to leverage brain-sensing technologies to build cognitive neural prostheses or brain-computer interfaces (BCIs), in which users explicitly manipulate their brain activity instead of using motor movements to control computers. For example, paralyzed patients can now control a cursor, type text, or move a wheelchair simply by imagining moving different parts of their bodies or by thinking about different tasks. The next generation of devices promises higher bandwidth, potentially allowing direct control over devices such as robotic limbs.

Regardless of the actual method, successful input typically requires adequate and appropriate system feedback to guide actions, confirm actuation, and present results. This feedback, or output, is presented through a form that is perceived by the human. The most common form of output has been visual output through computer displays, and the subfield of information visualization has focused on exploiting principles of human perception and cognition to design imagery that best conveys ideas. In addition to visual output, designers have also explored the use of auditory, tactile or touch, and even olfactory and gustation interfaces to take advantage of the broad range of human sensory channels. One example of compelling tactile output is game console controllers that vibrate when hit by an opponent. Similarly, many global positioning systems (GPS) units now have auditory interfaces in addition to the traditionally visual map because drivers cannot always free their eyes from the task at hand to attend to information.

Evolution

The evolution of the human–machine interface can be divided into several historical phases, marked by the dominant interface of the time. In the early days, the prevalent model was batch interfaces, in which users specified all details of a task—for example, on physical punch cards—executed them, and received results only when the processing was fully completed. This was tedious and error prone and was followed by developments in command-line interfaces, which allowed users to interactively issue commands the system immediately executed and provided results for. Although an improvement, command-line interfaces did not take full advantage of human perceptual, cognitive, and learning abilities, which were leveraged with the development of graphical user interfaces (GUIs). In modern GUIs, users engage in rich communication with the computer using the various input devices described earlier. For example, in the window, icon, menus, pointer (WIMP) model or the desktop metaphor, users manipulate virtual objects on screen as if they were physical objects such as files and folders on a desk or a trash can on the floor. Researchers are now working to invent the next large human–machine interface paradigm shift, commonly referred to as post-WIMP interfaces.

Desney S. Tan

See also Action and Vision; Computer Graphics and Perception; Computer Speech Perception; Multimodal Interactions: Visual–Haptic; Neural Prosthetic Systems

Further Readings

Baecker, R., Grudin, J., Buxton, W., & Greenberg, S. (1995). *Readings in human–computer interaction: Toward the year 2000* (2nd ed.). San Francisco: Morgan Kaufmann.

Card, S. K., Mackinlay, J., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think*. San Francisco: Morgan Kaufmann.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human–computer interaction.* Hillsdale, NJ: Lawrence Erlbaum.

Lindsay, P. H., & Norman, D. A. (1977). *Human information processing: An introduction to psychology* (2nd ed.). New York: Academic Press.

I

Imagery and Perception

See Auditory Imagery; Olfactory Imagery; Visual Imagery

Impossible Figures

When we look at a picture, we can often see the two-dimensional lines and colors on the surface that comprise the picture and what the picture depicts—what it is a picture of—which is typically a three-dimensional object such as a person or a landscape. Depiction is a special kind of pictorial representation, which many think is mediated by resemblance between the picture and what is depicted (although this is contested by some philosophers). An example of non-depictive pictorial representation is the traditional representation of saints. This does not rely on capturing how the saint looked (for that is often unknown) but, rather, employs a symbol system, wherein the objects near to or carried by the saint determine the saint's identity.

Many pictures depict things that exist and many depict things that don't. For example, a picture might depict Glasgow (which exists) or Brigadoon (which doesn't). A few special pictures seem to depict things that could not exist. These pictures are referred to as "impossible figures" and the objects they depict are "impossible objects." A characteristic of these pictures is that it may take an observer some time to realize that the figures are impossible. Artists, psychologists, mathematicians, computer scientists, and philosophers have studied impossible figures, as discussed in this entry.

Examples

Consider first, perhaps the most well-known impossible figure, the tri-bar, a version of which was first drawn in 1934 by the Swedish artist Oscar Reutersvärd (see Figure 1a). Roger and Lionel Penrose subsequently discussed the figure in print in 1958. The picture is naturally seen as depicting a three-dimensional object consisting of three bars joined at the ends to form a closed figure. But we can easily recognize that such an object is impossible. Consider the mutually orthogonal Cartesian axes x (horizontal), y (vertical) and z (depth). With respect to distance from the origin along the z-axis: Corner c is further than corner a, corner a is further than corner b, and corner b is the same distance as corner c. This entails that corner c is further from the origin along the z-axis than itself, which is logically and geometrically impossible.

This kind of impossibility, where points on an object have an inconsistent position in space, is manifest in many different kinds of impossible figure. For example, impossible bar figures with a greater number of bars than the tri-bar can easily be drawn (e.g., Figure 1b), and they all exhibit this kind of impossible nature, as do impossible three-dimensional cubes (Figure 1c). Other figures that display this type of inconsistency but along the y-axis, are the impossible staircase (Figure 1d) and Turner's ziggurat (Figure 1e).

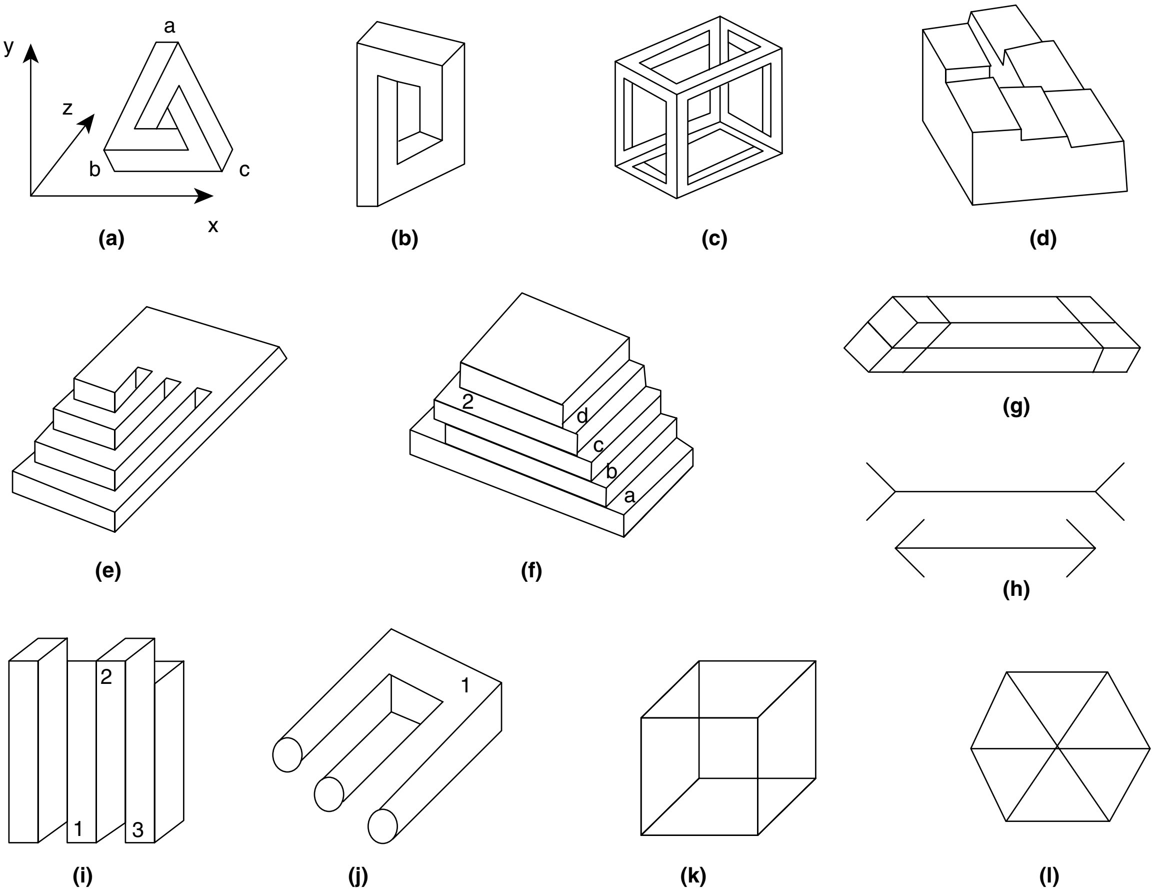


Figure 1 A Selection of Impossible Figures, Ambiguous Figures, and Illusions

A special kind of inconsistent spatial position is exhibited by impossible objects whose planes seem to have two different orientations at the same time. In Figure 1(f), the horizontal part of the step labeled 1 turns into the vertical plane between the horizontal parts of the steps labeled a and b. Similar inconsistencies apply to the other steps. Figure 1(g) can be seen, on the left side, to consist of two blocks stacked one on top of the other and, on the right side, to consist of two blocks side-by-side, yet they seem connected by straight bars.

A different kind of impossible object violates rules that govern the boundaries of solid objects. A part of the figure that appears to depict part of a solid object also seems to depict the surrounding air. Parts of Figure 1(i) look like parts of solid blocks, labeled 1 and 2, but they fizzle out into thin air. Figure 1(j), known as the devil's tuning fork, seems to contain a solid area, labeled 1, but that area dissolves into the surrounding space. (These two figures also contain areas of spatial position inconsistency.)

Disciplinary Interest

Art

These figures have inspired numerous artists, among whom are the aforementioned Reutersvärd and, notably, M. C. Escher. Unlike the bare outline drawings in Figure 1, these artists often use subtle cues of perspective and shading that heighten both the seeming reality of the objects and their impossibility.

Mathematics

Mathematicians have tried to classify certain impossible objects according to different principles and tried to find algorithms for determining whether a depicted object is impossible. Different approaches have met with varying degrees of success and criticism. Such algorithms would be helpful in developing computer vision, enabling computers to reject interpretations of the world that yielded impossible objects.

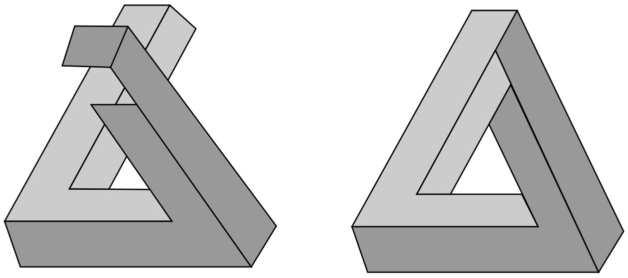


Figure 2 Pictures of a Possible 3-D Object as It Would Look From Different Angles

Psychology

Psychologists have been perhaps most interested in impossible figures. A question that perception of impossible figures raises is this: Why do we continue to see these figures as representing impossible objects when we realize that they are impossible? Why don't we see the figures as two-dimensional lines on the page or as representing other three-dimensional possible objects that are consistent with the lines on the page? Some figures, such as Figure1(c), initially look just like two-dimensional lines on the page, but with practice, these can be seen as three-dimensional objects, in this case, a cube. Note that all pictures are compatible with an infinite number of real-world scenes, only one of which, or sometimes in the case of ambiguous figures such as Figure 1(k), a low number of which, we take the picture to depict. (Figure 1k can be seen as a cube facing down and to the left or facing up and to the right.)

Part of the answer is that the visual system is, at least to some degree, autonomous from our belief system. For example, visual experiences of illusions persist, even when one truly believes that one is undergoing an illusion. For example, the horizontal lines in the Müller-Lyer illusion, Figure 1(h), continue to look to have different lengths, even when one knows that they do not. But this is not the whole answer. For the question remains: Why does the visual system not autonomously alter the interpretation?

Many psychologists think that the answer is that the visual system works by following certain rules relatively rigidly, even when doing so yields an inconsistency. By studying impossible figures (and other perceptual phenomena), psychologists try to work out these rules. For example, when looking at the impossible tri-bar, they believe that the visual system follows the law of proximity. This states that if there is no visible gap between elements depicted in a picture, then they are perceived as forming a totality and joining up. This is why the figure is not seen as an object with shape as in Figure 2(a), which lacks conjoined ends, viewed from an angle where there was no visible gap between the ends.

The visual system follows such rules all the time, not just when looking at pictures. Psychologist Richard Gregory constructed an object out of wood with the shape of Figure 2(a). He demonstrated that viewing it from a particular angle made it look like an impossible object existing in the real world (that depicted in Figure 2b). He photographed the object from that angle producing a photograph that looked like an impossible object.

Psychologists have also studied whether different subsets of the human population react differently to impossible figures and, if so, what this shows about their minds. For example, children can realize that some but not all impossible figures depict an impossibility. Results support the idea that the children see depth and solidity in the figures. Children don't realize some figures are impossible because they cannot construct an internal model of them. Another example is people with autism and Asperger syndrome who have difficulty detecting that some pictures depict impossible objects. The reason for this is still under investigation but some evidence suggests these people focus on fine detail and local structure at the expense of overall context or global structure.

Psychologists have also found factors that affect how quickly normal subjects realize that a figure is impossible. Julian Hochberg reports that figures of the type shown in Figure 1(b), whose four angles are close together in space, are immediately recognized by normal adults to be impossible, but figures of the same type, whose four angles are far apart from each other, are not initially noticed to be impossible. This suggests that to realize that some figures are impossible, subjects have to scan various parts of the figure and integrate information from various local depth cues.

Philosophy

Philosophers are interested in impossible figures as a way of studying the content of visual experience. Visual experiences seem to represent objects and properties around us—such as shapes and colors—and these are the contents of the experience. Impossible figures (and Gregory's illusion of an impossible object discussed previously) provide prima facie examples of visual experiences with explicitly contradictory content—such as the content that a plane is horizontal and that it is not horizontal (for example, Figures 1f and 1g). (This is significant because many philosophers hold that beliefs cannot have explicitly contradictory content.)

Philosophers are also interested in the nature of pictorial depiction. One question, already noted, is whether depiction involves resemblance, or experienced resemblance, between a picture and what it depicts. Depiction of impossible figures is an interesting case because the figures don't and couldn't exist, so how could a picture resemble or be experienced as resembling them?

Another philosophical issue concerns whether, apart from the geometrical logical impossibilities depicted in the impossible figures discussed thus far, other (non-geometrical) logical impossibilities can be depicted. If not, why not? If so, why don't we have a clear example?

Finally, an important issue arises when we consider what kinds of geometrically impossible objects it is possible to depict. It does not seem possible to depict all kinds of geometrical impossibilities in a picture. For example, it seems impossible to depict a round square. Why can't we depict this kind of impossibility yet can depict the tri-bar? A likely answer is that the tri-bar is divisible into parts, each of which is perfectly possible. What is impossible is the *combination* of the parts as depicted. A good hypothesis is that all the objects depicted by (geometrically) impossible figures are divisible into parts, each of which is possible and each of which can be unproblematically depicted. (A similar hypothesis is that all pictures that depict impossible objects are divisible into parts, each of which still depicts a part of the original object, and each of which depicts a part of an object that could exist.) A round square can't be depicted because each part of it would have to be depicted as impossible—each part would have to be drawn as both straight and curved at the same time. Impossible objects that cannot be decomposed into possible parts cannot be depicted.

Fiona Macpherson

See also Bistable Perception; Computer Vision; Depth Perception in Pictures/Film; Infant Perception; Object Perception; Philosophical Approaches; Pictorial Depiction and Perception; Theoretical Approaches; Visual Illusions

Further Readings

Ernst, B. (1986). *Adventures with impossible figures.* St. Albans, UK: Tarquin.

Gregory, R. L., & Gombrich, E. H. (1973). *Illusion in nature and art.* London: Duckworth.

Kulpa, Z. (1987). Putting order in the impossible. *Perception, 16,* 201–214.

Penrose, L. S., & Penrose, R. (1958). Impossible objects: A special type of illusion. *British Journal of Psychology, 49,* 31–33.

Sorensen, R. (2002). The art of the impossible. In T. Gendler & J. Hawthorne (Eds.), *Conceivability and possibility.* Oxford, UK: Oxford University Press.

Young, A. W., & Deregowski, J. B. (1981). Learning to see the impossible. *Perception, 10,* 91–105.

INDIRECT NATURE OF PERCEPTION

Our everyday experiences attest to the immediacy of our perceptions. We need only open our eyes to see what is before us, much as the simplest of cameras needs only to be activated to capture a picture of what is visible through its lens. All that seems to be necessary in either case appears to be a lens that focuses the incident light rays and a light-sensitive surface upon which the rays can be focused, thereby creating an image of the object(s) reflecting the light. In the case of a camera, the film, or film equivalent, is the light-sensitive surface, but in the eye, the retina serves this function. Despite the appearance of immediacy, however, our perceptions are not immediate nor are they simply pictures in our heads.

Two main theory types, *direct* and *indirect,* have been proposed to account for what goes on between the registration of the light rays on our retinas and our perceptions. In general, these two types of explanations differ primarily in their reliance on what is referred to as *extra-retinal information*—information not contained in the retinal image. This entry discusses indirect theories of perception, in contrast with direct theories, and provides additional characteristics common to all indirect theories.

Direct and Indirect Theories of Perception

First, let us consider a typical example of a perception that illustrates the claim that perceptions are not pictures in our heads, a claim shared by both direct and indirect types of perceptual explanations, and that reveals the sharp difference between how these two theory types account for what we see. Consider, for example, that stationary objects in our line of sight continue to appear stationary to us even though their retinal images move constantly because they move every time our eyes or heads move, and our eyes, at least, move constantly. This is a phenomenon known as *position constancy* or the perception of stability. Given this, if perceptions were pictures in our heads derived directly and solely from the images on our retinas, it would follow that we would never perceive objects in the world as stationary. These retinal image movements produced by our own movements generally do not lead us to see stationary objects move, so there must be more here than meets the eye. That "more" is the difference between indirect and direct accounts of perception. Before turning to how these two theory types account for position constancy that will more concretely illustrate the differences between them, this entry presents a brief summary of how these theories, which have been around in one or another version for more than 100 years, differ generally.

Direct Theory

The defining difference between these two theories is centered on the nature of the information deemed necessary for perception to occur. In general, direct theories of perception, the first version of which is attributed to Ewald Herring (1834–1918) and the most recent attributed to J. J. Gibson, claim that our perceptions can be fully explained by the information available in the optic (retinal) array—that is, the patterns of light reaching the eye that, according to this theory, provide unambiguous information about the layout of objects in space. According to Gibson, our perceptual system is designed to pick up *invariant information,* that is, aspects of the environment that do not change over time despite our movements or the movements of objects. For example, as we move forward in our environments, the objects that are straight ahead of us remain centered and stable on our retinas, but all the other objects in our environment move outward in a sunburst-like pattern invariably indicating what we are moving toward. This does not mean, however, that theories of this sort deny that considerable internal processing occurs before the act of perception—if so, they would be proposing a version of the picture in the head theory of perception. Rather, what these theorists propose is that whatever these internal processes are—which on one view involve an internal resonating, or tuning fork-like process that detects invariant aspects of the optic array—they *do not* rely on other nonretinal information. For example, a direct theory of perception denies that these internal processes are in any way cognitive-like or dependent on our past experiences.

Indirect Theory

In contrast, all indirect theories, the early version of which is attributed to Hermann von Helmholtz (1821–1894), claim that the information available in the retinal image array is *insufficient* and therefore cannot account for our perceptions without using additional information not available in the image and without the operation of cognitive-like processes. Thus, like direct theorists, these theorists also propose a set of internal processes that intervene between image registration and perception, but these processes are markedly different from those proposed by the direct theorists. The internal processes relied on by indirect theorists avail themselves of our experiences, our expectations, or our implicit hypotheses about the nature of the visual world and are ratiocinative or cognitive-like in nature. For example, they have been described as inferential (Helmholtz), as a form of hypothesis testing (Richard Gregory, Julian Hochberg), or as rational (Irwin Rock). Another way in which these processes are said to differ is that they operate from the top-down—that is, from the cortex down to the sensory processes—whereas those proposed by the direct theorists are considered bottom-up and thus operate in the opposite direction.

Position Constancy

Direct Theory

These differences are apparent in the ways in which the direct and indirect account for perceived stability. Direct theories of perception account for our perception of object stability in the face of image movements caused by our own movements strictly in terms of the behavior of images on the retina when we move. The theory points to the fact that when we move in a stationary environment, for example, all the retinal images of all the objects we can see move homogeneously on our retinas; that is, they all move in the same direction and at the same rate because their motion is solely a function of our own motion. What is important is that there is no relative motion among any of the images, whereas when an object moves in our visual environment, or when a person walks across our field of view, its motion is different from the motion of all the other images of stationary objects because it is a joint function of the object's own physical motion and our motion. Thus, when an object is moving in the environment, there is relative motion between the image of the moving object and all the other images on our retinas that are derived from stationary objects, and when we move our eyes when viewing a stationary environment, all objects in the environment sweep across our retina together (or must move together) without any relative movement between them.

Thus, homogeneous image motion is an invariant characteristic of image motion caused by our own movements, and when it is detected by the internal processes attuned to picking up invariance, we perceive objects as stationary. When images move relative to one another, that is, move in different directions and at different speeds (something that invariably occurs when an object we are looking at moves), then we perceive its movement. Relative image motion is the invariant characteristic of image behavior associated with the object motion, whereas homogeneous image motion signifies object stability. In either case, our perceptions of both stability and motion can be, according to this theory, fully explained by information derived entirely from the optical array.

Indirect Theory

In stark contrast, indirect theories account for perceived stability in the face of image motion caused by our own movements in terms of extraretinal information—that is, information not available in the optic image but rather derived from other sources of information, in this case information about our eye and head movements. One version of an indirect account of position constancy proposes a hypothetical process, the function of which is to compare information about image motions that is derived from the retina with information about eye and head movement derived from head and eye motion signals. When these two sources of information match, no object motion is perceived and image motion is attributed by the perceptual system to the movements of the observer. When they do not match, an object is perceived to move in the direction and to the extent specified by the mismatch. For example, if we move our heads to the left, images in our visual environment displace to the right by the same amount and at the same rate, thus canceling each other out and signifying that the movement in our visual field is caused by our own motion. Conversely, if we move our heads to the left and the visual field does not displace, this signifies that the objects in our field are moving at the same rate and in the same direction as our heads. Indirect theorists argue that their explanation in this instance and in many others can account for examples of perceived motion and stability that direct theories cannot explain.

To cite one example, if there is only a single object in our visual scene that is moving and that we follow with our eyes, its image will not move on the retina (because it is as if it were glued to our retina); nevertheless, we perceive that object as moving. For the indirect theory, this follows because there is a mismatch between the information that the eye is moving and the information that the retinal image of the object is stationary. It therefore follows that the object must be perceived to move, and it is. However, if relative image motion is the defining characteristic of object motion according to a direct theory of perceived motion and stability, then the object should not be perceived to move because in the absence of any background, there is no relative retinal image motion—in fact, there is no image motion at all—and yet the object is perceived to move. Another well-known phenomenon where there is also no retinal image motion but we have a vivid sense of object motion is the case of stroboscopic motion, which in its simplest form occurs when two points of light separated in space flash on and off in sequence and we perceive a single object moving back and forth. This is another case that is not easily explained by a direct theory of the perception of motion because there is no relative image motion to account for our perceived motion.

Additional Distinguishing Characteristics of Indirect Theories

A number of other characteristics are shared by all theories of indirect perception. Foremost is the assumption that the information in the visual stimulus—that is, the retinal array—is insufficient and does not unambiguously specify a particular percept, as we have seen in the example of perceived stability described earlier. To briefly describe another example, consider the case of perceived size constancy, which refers to the perceptual fact that our perceptions of the size of an object remain constant despite huge variations in the size of its projected retinal image caused by the change in the distance between ourselves and the object. That is, even when 6-foot-tall persons double their distance from us, which halves the sizes of their images on our retina, we still perceive the persons to be 6 feet tall. An indirect account of size perception entails two kinds of information: information about image size, which is directly available, and information about the object's distance, which frequently may not be. In other words, the perceived size of an object is derived or *inferred,* according to von Helmholtz, from information about its distance. The object's image size is smaller, but this is compensated for by the increased distance. There is an opposing direct account of perceived size constancy, but it has great difficulty accounting for our ability to perceive the size of an object when it is the only thing we can see.

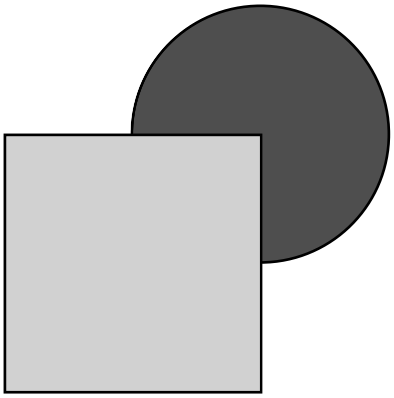


Figure 1 The Principle of the Avoidance of Coincidence

Notes: It appears to us as a square in front of, and occluding a disc, when what is actually pictured is a disc with a piece missing aligned with the square. Our perceptual system rejects the latter interpretation of the input as too unlikely a coincidence.

Rules of Inference

Although the main issue of contention between direct and indirect accounts of perception concerns whether information available in the retinal array is sufficient to account for our perceptions, this is not the only issue that differentiates the theories, although some of the other differences are not characteristic of all indirect theories. One that does, however, characterize all indirect theories is the conviction that other universal principles or rules govern the perceptual processes that act on, disambiguate, or make sense of retinal stimulation, although the theories differ about what these rules may be. On one account, the processes obey a *likelihood* principle, that is, given a set of possible scenarios that might have caused a particular retinal array, the perceptual processes opt for the one that is most likely where likelihood might be based on past experience or some other rubric. A related proposal, by Rock, is that perceptual processes obey the principle of the *avoidance of coincidence,* so that what is perceived is not, if possible, an unlikely coincidence. See, for example, Figure 1, which appears to us as a square in front of, and occluding a disc when what is actually pictured is a disc with a part removed aligned with a corner of the square. Our perceptual system rejects the latter interpretation of the input as too unlikely a coincidence. So, just as cognitive processes are rule governed, so, according to these theorists, are the inferential processes at work in the perceptual system. These rules all shape the cognitive-like processes, the purpose of which is to figure out what visual scene produced the momentary set of retinal images.

Arien Mack

See also Direct Perception; Ecological Approach; Theoretical Approaches; Unconscious Processes

Further Readings

Epstein, W. (1973). The process of "taking into account" in visual perception. *Perception, 2,* 267–285.

Gibson, J. J. (1966). The theory of information pickup. In *The senses considered as perceptual systems* (pp. 266–286). Boston: Houghton Mifflin.

Rock, I. (1997). *Indirect perception.* Cambridge: MIT Press.

Rock, I., & Brosgole, L. (1964). Grouping based on phenomenal proximity, *Journal of Experimental Psychology, 67,* 531–538.

INDIVIDUAL DIFFERENCES IN PERCEPTION

Our individual views of the universe may be different from one another's because we each encounter only one small part of what is there. The ancient Hindu parable of the six blind men and the elephant—wherein each man describes only the part of the elephant he is touching, forming an incomplete representation of the whole—is an illustration of such individual differences. The elephant, a metaphor for the universe, is perceived by one man as a snake (because he feels the trunk), by another as a tree (because he feels the leg), and so on. However, individual differences in perception are not usually as large as the differences between the six blind men's percepts of the elephant. The differences are not so large because perception is a complex phenomenon resulting from multiple small effects, such as many different genes and accumulated experiences, acting mostly separately.

When a large number of small effects act additively and independently, the resulting observations will be normally distributed. Sir Francis Galton coined the term *normal distribution* for the bell-shaped curve that describes how variables such as height in inches are distributed in the human population. In his 1869 book *Hereditary Genius,* he suggested that mental faculties such as perception are distributed in a similar fashion. Figure 1 shows gender-related differences in performance on a line-angle matching task from 250,000 respondents to an online quiz (www.bbc.co.uk/science/humanbody/sex). Men were better than women at doing this task, but the difference between the most successful and the least successful woman was 10 times larger than the average difference between men and women, and most people of both genders fell somewhere in the middle. This means that of three random male-female pairs, most likely in at least one the woman will outperform the man. Concepts such as "femininity" and "masculinity" must not be used as strict categories because single variables such as gender, though they help shape society, are of little importance in determining how a given individual will perceive the world. This entry discusses various factors that lead to individual differences in perception.

Gender-Related Differences

The worlds perceived by men and women are subtly different. On average, men outperform women on tasks of visual-spatial discrimination, such as matching the angle of a line, whereas women frequently outperform men on visual tasks involving color discrimination and discerning facial expressions. Hearing is significantly more sensitive in newborn female infants, compared with male infants. Women of all ages hear better than men do above 1,000 hertz (Hz), but men outperform women at 500 Hz. Hearing impairments resulting from age are twice as prevalent in men as they are in women. Women are generally more sensitive to

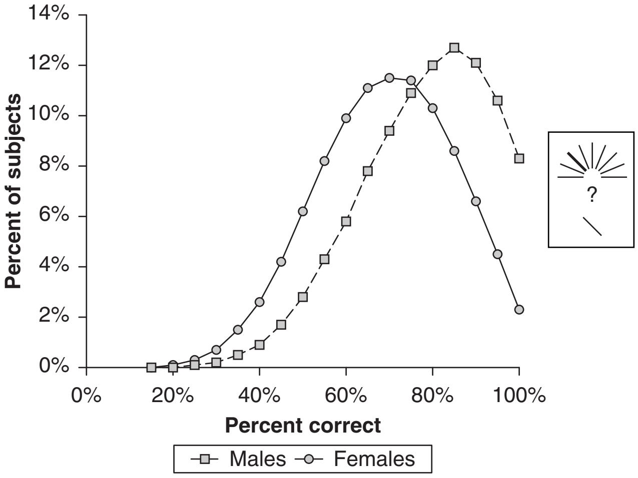


Figure 1 Gender-Related Differences in Perception

Source: Collaer, Reimers, & Manning (2007).

Note: Task was to match a comparison line to one of the presentations in a fan display (see box).

odors than men are, especially in youth and old age. Women and men vary in their sensitivity to pain. Women are more likely to find a given stimulus painful but are better able to discriminate between different levels of pain.

Age-Related Differences

Human infants are born with a slightly limited capacity to make perceptual discriminations, and this capacity improves through learning and maturation until the age of about 10 to 20 years. Thereafter, sensory discrimination declines, with catastrophic sensory losses beyond the age of about 80 years. Through a healthy life span, the perceptual world begins and ends as slightly fuzzy and indistinct, and people much older or much younger than ourselves are not able to discriminate shapes, colors, sounds, smells, flavors, and bodily sensations as well as we can. For example, if you are 60, you can probably decipher smaller letters than the average 90-year old, or the average newborn infant.

Mean visual acuity varies from about the third line of a standard optometrist's chart in infancy, to the smallest line by age 10, and back to the third line by age 80 years and older. Beginning around age 40, the lens inside the eye gradually stiffens, which impairs the ability to focus the image on the retina by changing the thickness of the lens (known as accommodation), eventually making corrective lenses necessary for near vision in most individuals. Changes in the color of the lens are the cause of some age-related alterations to color vision. Color vision is most precise between the ages of 20 and 30, when people can accurately sort at least 50 different shades and hues.

The ability to hear high-pitched sounds deteriorates with age, to the degree that a cell-phone ringtone about 3 octaves higher than the highest note on a piano was marketed to teenagers on the Internet for its use in the classroom, undetected by teachers. Age-related hearing loss within the pitch range of human speech occurs in one of every three adults aged 65 to 75, and in twice as many men as women.

Odor discrimination peaks between 20 and 50 years of age. More than half those aged between 65 and 80 years, and three quarters of those older than 80, experience significant impairments in odor discrimination.

It was once thought that newborn infants lacked the ability to perceive pain because the neural apparatus for its perception was undeveloped, but this is now known to be false. As with the other senses, pain sensitivity increases through childhood, and then diminishes in adulthood, although these changes are relatively small compared with the variability in pain perception between different individuals.

Differences Due to Experience

A babbling infant quickly learns to home in on the sounds that are relevant for the creation of meaning in his or her language environment. For example, Japanese speakers have great difficulty distinguishing the phonemes /r/ and /l/ in English because both sounds have the same meaning when used in words in their own language. Such skills are generally acquired early in life. The connections between brain cells that determine how we perceive, think, and act are mostly set in place during childhood, but they can undergo significant changes at any age. This is called *experience-dependent plasticity*. For example, video-game players have double the chance of detecting small targets in the periphery, or attending to a rapid sequence of targets in the center of the screen, compared with non-video-game players. Games such as Nintendo's Brain Age that engages the player in puzzles such as Sudoku capitalize on such evidence. A recent study showed that adults older than 65 improved at hearing and remembering speech sounds after intensive training. Professional wine tasters outperform others on tests of odor discrimination. Musicians outperform nonmusicians on pitch discrimination in certain chords and may have an enlarged region of the brain devoted to auditory processing. This kind of learning takes years of practice, but most of us are familiar with the experience of our own perception changing in an instant, as when we first learn to read a word and can never see it again without understanding its meaning. Perceptions differ sharply depending on what we have experienced in the past, and individuals who have practiced a particular task perceive it differently. One's perception can change at any age as a result of experience-dependent plasticity.

Perceptual Deficits and Peculiarities

Sensory impairments are a major cause of individual differences in perception. Common perceptual and sensory deficits affect large numbers of individuals. Color deficiency affects about 7% of the U.S. male population, but less than 1% of females. These individuals generally have difficulty discriminating red and green. This kind of deficiency is caused by genetic differences. Temporary disturbances of sensation, such as tunnel vision, can result from cardiac events and occur in about 4% of adults older than age 65. Blindness and low vision have a prevalence of about 3% in U.S. adults older than 40. About 16% of the adult population (ages 17–80 years) suffer from significant hearing impairment, based on data from the United Kingdom.

As well as not perceiving what there is, many people perceive what is not there. Hallucinations are a common experience in psychiatric patients, and in the general population. A survey of more than 13,000 people in Europe showed that almost 40% of people reported some hallucinatory experiences. The most common forms were infrequent (less than once per month) hallucinations of touch in the moments before sleep, that occurred in 16.9% of the people surveyed, and daytime odor hallucinations, occurring in 8.6% of people. Daytime visual and auditory hallucinations were most commonly associated with mental illness, whereas daytime touch hallucinations were associated with current use of drugs. This means that for any random two people, most likely at least one of them sometimes feels, smells, sees or hears things that are not there.

Synesthesia is a rare condition in which one sense is confused with another. For example, Galton wrote in 1880 about some people, mostly women, who confused numbers with colors. This led to an x-linked dominant gene hypothesis, which was subsequently disproved by the findings that identical twins can be discordant for synesthesia, synesthesia can skip generations, and population-based studies show no gender bias. Different people volunteer that they confuse their senses from the people who admit they do when asked, so research methodology may be responsible for conflicting findings. Whole edifices of genetic theory can be built on such tainted findings. Culture influences perception as well as our perception of culture influences our actions.

Culture and Perception

Just as /r/ and /l/ may sound the same to Japanese speakers, so color differences may vary in their noteworthiness depending on the language culture. Brent Berlin and Paul Kay wrote a book in 1969 about how human cultures learn to name different colors in a certain order through their evolution, just as infants learn to name colors during development. In color, as with the chemical senses, genetic differences are notable. Nevertheless, by associating color names with things, we can reach consensus on what they mean. Although you and I may have different proportions of red, green, and blue cones in our eyes, we can both agree that green is the color of grass leaves. By referring to things outside of our own perceptual systems, culture helps us to communicate and make meaning of our senses.

Genetic Differences

Genetic differences offer a handle on some of the many causes that contribute to the normal distribution of perceptions in the human population. Scientists can now, for the first time, tell the story of how certain genes are expressed in the body as proteins that transform physical energy from the sensory environment into neural impulses that determine what we perceive. Individual differences in perception are quintessential to this enterprise because they allow scientists to identify the genes that cause some people to have different photopigments for color, or different taste receptors, by looking at what colors and flavors we each perceive. Scientists predicted the severity of color vision loss in 16 individuals with color abnormalities by looking at their photopigment genes. Naturally occurring gene-alleles in the human population determine sensitivity to certain bitter and sweet molecules. Not everybody can detect the same flavors, and perhaps this is one reason why we have different tastes.

Philosophical Implications

The 18th-century philosopher Immanuel Kant said that sensory experience tells us what is there, but it doesn't tell us which of the things that we perceive are truly fundamental facets of the world rather than the chance effects of the ways in which our sensory systems function. He claimed that our ability to reason gives us a desire to know the true nature of the world out there. If we put many individuals' experiences together (weighting them by their precision for a given judgment, so that, for example, we defer to youth in hearing high-pitched sounds and defer to young adults in the perception of color) then we can arrive at a more faithful representation of the world, just as the six blind men in the Hindu parable might combine their impressions of the elephant into a more accurate model of the whole than any individual alone, whereas if they all felt the same part of the elephant they would agree with one another, but their ability to access universal knowledge of the elephant would be compromised. Our individual differences in perception, from sensory perception to more complex social and moral perceptions, provide the whole of society with more knowledge than could be gleaned if we all perceived in exactly the same way. It is important to remember and respect these differences, and not to assume that one arbitrary individual's perception can usefully serve as an exemplar for the whole of humanity because it might lie anywhere under the normal distribution for the property being measured. However, just as the blind men will never perceive the elephant's grey color, so there are likely to exist properties of the physical and social worlds that by virtue of our conjoint limitations and despite our differing perspectives, we will never experience or know. The worlds we all perceive are more similar than differing.

Ariella Popple

See also Ageing and Chemical Senses; Ageing and Hearing; Ageing and Touch; Ageing and Vision; Cultural Effects on Visual Perception; Emotional Influences on Perception; Neuropsychology of Perception

Further Readings

Collaer, M. L., Reimers, S., & Manning, J. T. (2007). Visuospatial performance on an Internet line judgment task and potential hormonal markers: Sex, sexual orientation, and 2D:4D. *Archives of Sexual Behavior, 36,* 177–192.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423,* 534–537.

Hubbard, E. M., & Ramachandran, V. S. (2005). Neurocognitive mechanisms of synesthesia. *Neuron, 48,* 509–520.

Neitz, J., Neitz, M., & Kainz, P. M. (1996). Visual pigment gene structure and the severity of color vision defects. *Science, 274,* 801–804.

Ohayon, M. (2000). Prevalence of hallucinations and their pathological associations in the general population. *Psychiatry Research, 97,* 153–164.

Vitello, P. (2006, June 12). A ring tone meant to fall on deaf ears. *The New York Times.*

Walco, G. A., & Harkins, S. W. (1999). Lifespan developmental approaches to pain. In R. J. Gatchel & C. Dennis (Eds.), *Psychosocial factors in pain: Critical perspectives.* New York: Guildford Press.

INFANT PERCEPTION

Late in the 19th century, the pioneering psychologist William James in *The Principles of Psychology* famously summarized what he believed to be the infant's perceptual experience, "The baby, assailed by eyes, ears, nose, skin, and entrails at once, feels it all as one great blooming, buzzing confusion," and suggested, "Infants must go through a long education of eye and ear before they can perceive the realities which adults perceive. *Every perception is an acquired perception*" (p. 488, emphasis original). This position was echoed by the preeminent developmental psychologist of the 20th century, Jean Piaget, who proposed that at birth, percepts across sensory modalities, such as vision and touch, are uncoordinated and that the newborn's visual world consists of an assemblage of moving colors and shapes, rather than the segregated objects experienced by adults. Perceptual organization emerges only slowly during the first two postnatal years, according to Piaget, as infants gain direct manual experience with objects and the coordination of visual, auditory, and tactile information.

Systematic investigations of infant perception were not reported until about 50 years ago, with the pioneering work of Robert Fantz. Part of this delay (relative to studies of adult perception) was caused by limitations in methodology. Fantz discovered that infants would often show visual preferences, looking longer at some patterns relative to others when the stimuli were presented sequentially. This allowed scientists to begin to chart changes in visual capacity with development. A second important contribution made by Fantz was the finding that repeated presentations of stimuli often resulted in a decrement of attention in infants. Subsequent research refined these methods and eventually led to the adoption of more sophisticated paradigms such as the habituation/dishabituation method. This and related designs capitalize on the tendency of infants to look longer at stimuli that are unusual or unexpected, and these methods have revealed a wealth of cognitive and perceptual competencies in infancy. The methodological repertoire of infant researchers in the past few decades has expanded considerably and now includes precise recordings of eye movements, brain activity, conditioning of behaviors such as head turning or reaching, and other behaviors under both controlled and naturalistic conditions.

The results of experiments that use these methods have revealed that all sensory systems in infants are functional at birth, and even coordinated to some degree. These investigations help us understand both how infants experience the world and how their knowledge develops over time as a function of the interplay of existing and developing sensory systems, actions, and physical development, including development of the brain. At birth, vision is somewhat organized, and audition, olfaction, and touch are fairly mature. Intermodal perception, coordination of information from a single event through multiple modalities, begins early, and improves across infancy.

Visual Perception

Basic Visual Function

Vision is arguably the most important sensory modality available to humans because it provides the perceiver with essential information about near and distant objects and events so that appropriate actions can be planned. Most basic visual functions are operational yet relatively immature at birth. Visual acuity, the ability to distinguish fine detail or differences in adjacent positions, is estimated at 20/200 to 20/400 for most newborns. Acuity improves rapidly within the first few months. Contrast sensitivity, the ability to detect luminance differences between two adjacent areas such as stripes on a grating, is also reduced in newborns relative to adults and develops as infants gain visual experience. Color vision is not yet fully developed at birth, but infants are able to see colors nearly as well as adults can by age 4 to 6 months.

The ability to perceive motion is particularly helpful to infants as they begin to disambiguate the visual world. Objects and people in the environment move in many different ways (laterally, vertically, toward and away from the observer, and rotating) and at different velocities. Infants' responses to the slow and fast motion velocities differ depending on the types of motion and age, suggesting the existence of separate mechanisms for processing different types of motion. Moreover, infants' own motion also contributes to motion perception. Despite the complex nature of motion, however, nearly all types of motion perception develop by 6 months after birth.

In similar fashion, depth perception gradually develops during the first several months. Infants first become sensitive at about 2 months to kinematic, or motion-carried information for distance, as when one surface moves in front of another. Next, at about 4 months, infants are able to perceive depth via the difference in the optical projections at the two retinas to determine depth, known as stereopsis. Stereoscopic depth cues provide information about distances of objects in near space as a function of their relative horizontal positions in the visual field. Finally, at about 7 months, sensitivity to pictorial cues allows infants to perceive depth in a flat, two-dimensional picture.

Visual Attention

Infants are born with a functional oculomotor (eye movement) system. The muscles that move the eyes, and the brain-stem mechanisms that control these muscles directly, appear to be fully mature at birth, and infants make good use of these systems to scan the visual environment. Two developmental events seem to be particularly important to control of visual attention: the emergence of smooth pursuit at about 2 months, and increasing top-down control over saccadic or scanning eye movements that can take much longer. Smooth pursuit helps track moving targets in the environment and stabilize gaze, and saccades are used when inspecting visual stimuli. Both kinds of eye movement are believed to develop as specialized brain regions and networks mature, in particular systems for processing motion and objects.

Object Perception

Object perception is complex, involving multiple information-processing tasks, such as perceiving boundaries, shapes, sizes, and substances of objects. Understanding object boundaries first requires recognizing where one object ends and another object or surface begins. Detecting edges is critical for this process, and the intersection of edges provides information for the relative distance of object and surfaces. For example, where one edge is seen to lead into and end abruptly at another, the uninterrupted edge is usually nearer to the observer. Infants become capable of recognizing boundaries between 3 and 5 months.

Recognizing object boundaries alone does not reveal the complete shape of all objects because many objects are partly hidden by other surfaces nearer to the observer. Perception of partly occluded objects as complete is first accomplished at about 2 months, when infants view a partly occluded object moving horizontally behind an occluder. Common motion of visible parts of objects is critical for young infants' perception, although infants begin to perceive object unity in stationary displays at 6 to 7 months.

Finally, size and shape constancy supports an observer's perception of an object's true size and shape as remaining constant even when it is viewed at different distances and from different angles, which causes the size and shape within the visual field to change. Surprisingly, newborns, who have limited visual experience, appear to exhibit some sense of both size and shape constancy.

Face Perception

Studies with newborns have revealed preferences for facelike displays relative to other patterned stimuli, and consistent preferences for faces throughout infancy. Newborns' ability to recognize facelike stimuli suggests that infants may have a crude representation of faces before any experience with actual faces, or it may indicate that faces match inherent preferences for stimuli with similar spatial frequencies and arrangements of internal stimulus elements.

Under naturalistic conditions, faces exhibit multiple expressions and are seen from multiple perspectives, but infants are able to recognize familiar faces despite these variations. Infants can also discriminate gender in faces, and most infants show preferences for females. However, infants whose primary caregiver was male showed an overall preference for male faces, suggesting that experience with faces outside the lab leads infants to learn the details that define gender. Finally, infants' sensitivity to facial expressions emerges early; infants discriminate different intensities of smiling at 3 months and frowning at 6 months. By 7 months, infants can discriminate an extensive range of the facial expressions, including happiness, anger, sadness, fear, and surprise, although it is unlikely that they understand the content of this range of emotions as this age.

As noted, infants are predisposed to be sensitive to faces from birth. Some researchers have argued that this predisposition is caused by innate representation for faces, and others have suggested that it is a product of a combination of visual biases and rapid learning processes. Advances in neuroimaging and electrophysiological techniques have allowed researchers to identify areas in the brain involved in face perception. The middle fusiform gyrus in the right hemisphere has been implicated in perception of upright faces, and the amygdala appears to be involved with perception of facial expressions. Face processing in infants shows a right hemispheric advantage involving the fusiform gyrus, and experience with faces may help to sharpen the specializing of cortical areas tuned to faces. This tuning process may begin with a subcortical process involving the superior colliculus that orients infants to faces, manifest as the face preference described previously, and this can support learning about faces and the cortical sharpening process. Thus, infants' ability to perceive faces may begin as a response to a wide variety of facelike stimuli, including faces from other species; these abilities become tuned with age as a result of specific experiences with human faces.

Auditory Perception

Basic Auditory Function

Beginning with the second trimester after gestation, inner ears begin to function, which allows fetuses to have limited auditory experiences in the womb; as a consequence, fetuses show distinct responses to sounds of various intensities and frequencies. Evidence indicates that neonates' auditory perception is influenced by prenatal experiences with sounds. For example, newborns prefer listening to their own mother's voice to another woman's voice.

Despite the physical maturity of the cochlea about two-thirds of the way through gestation, sound conduction through the external and middle ear to the inner ear is inefficient at birth, hindering the transmission of information to the auditory neural pathway. Perception of low frequencies is poor in young infants relative to high frequencies; low frequency discrimination is not yet mature even at 10 years, but discrimination of high frequencies is superior in infants relative to that of adults.

The most common measure used when testing intensity processing is the absolute threshold, the smallest intensity of sound detectable in a quiet environment. The absolute threshold improves throughout infancy and reaches adult levels by puberty, and the higher the frequency, the earlier adult levels are achieved. For example, the absolute threshold level at 4,000 and 10,000 hertz (Hz) reaches adult levels by 5 years of age, whereas the level for 1,000 Hz requires 10 years or more to reach maturity. Between 1 and 3 months, the absolute threshold improves by 15 decibels (dB), but little improvement is observed between 3 and 6 months except another 15-dB improvement in the 4,000-Hz threshold.

Hearing in Naturalistic Environments

The studies on auditory discrimination summarized in the previous section were conducted with pure tones. Sounds in naturalistic environments are complex, comprising multiple frequencies and various intensities. For example, perception of timbre, such as hearing differences in the way different musical instruments sound, involves comparison of different intensities across different frequencies. As early as 7 months, infants can discriminate between sounds of different timbres with the same pitch, but adult levels of competence at discriminating a series of complex timbres are not reached well into childhood.

Localization, the ability to locate the source of sounds in space, is required to accurately perceive sounds occurring in naturalistic environments. Two processes are involved in localization: (a) evaluation of spectral shape and intensity and (b) binaural comparisons of sounds reaching the left and the right ears. Under normal circumstances, spectral shape is the primary cue to position in elevation, and binaural time and intensity differences are the primary cues to position in azimuth (the plane that runs through the ears parallel to the ground). Infants appear to rely more heavily on spectral shape in sound localization than on binaural differences, perhaps because they are more sensitive to differences in sound frequency than to differences in sound intensity.

Once different types of auditory information are received, they need to be organized into perceptually meaningful elements. For example, to follow a conversation, speech produced by members of the family must be grouped together and noises from children playing outside must be filtered out. The process of grouping is partly functional in infants, but it is more easily disrupted in children than in adults. Part of this process is ignoring irrelevant sounds while attending to the relevant sound source. Infants, unlike adults, often seem to act as if they are not sure about disregarding irrelevant sounds. For example, studies with 7- to 9-month-old infants suggest that they cannot detect a pure tone when presented simultaneously with a wide-frequency noise band.

Speech Perception

Infants appear to have difficulty segregating speech from other competing sounds. Thus, when interacting with infants, adult caregivers often compensate for this difficulty by making major acoustic adjustments in their speech, such as the use of infant-directed speech, which contains exaggerated pitch contours, a higher register, repetitions, and simpler sentences.

A central question in this area concerns whether infants respond to phonetic differences in a similar manner as adults. Studies examining cross-language and native-language speech perception suggest that infants are born with universal sensitivity to the phonemes that are present in all languages. Phonemes are components of a language that distinguish words by forming the contrasting element in pairs of words, such as the /r/ and /l/ in *rake* and *lake*. There is a developmental loss of “unused” initial sensitivities. For example, a study of English-speaking adults, Hindi-speaking adults, and 6- to 8-month-old infants from English-speaking families demonstrated that infants distinguished two distinct phonemes with similar sounds in both English and Hindi, /ta/ and /da/ in English, and the retroflex /D/ and dental /d/ in Hindi, whereas adults distinguished only between different phonemes in their native language. These phonemes are all produced by placing the tongue against the alveolar ridge, just behind the teeth, and releasing it in time with voice onset. They vary with respect to the precise part of the tongue and alveolar ridge involved, and voice onset timing; as mentioned, adults often have difficulty discriminating these phonetic details when they are not part of their native language.

Infants often exhibit preferences for speech over nonspeech sounds, which can help in attending to signals in the environment necessary for language acquisition. But infants do not always prefer speech. When presented with both filtered speech, which sounds like a kind of mumble as though hearing through a wall, and nonspeech sounds, neonates did not prefer listening to filtered speech, suggesting that preference for speech is not a direct result of prenatal auditory experiences with human speech. Instead, there may be an evolutionary predisposition for sounds produced by a human vocal tract. This claim has also been extended to a finding that hearing infants show preference for watching sign language over carefully matched nonlinguistic gestures.

The patterns of rhythm and intonation in speech are known as prosody. Newborns are sensitive to prosodic properties of speech and use them to discriminate one language from another; among several components of language, prosody appears to be the primary way for young infants to perceive speech. This is especially useful in bilingual environments because it helps infants avoid a potential confusion.

Intermodal Perception

Adults do not experience the world as fragmented sensory impressions, but as integrated, multimodal experiences. Infants are capable of perceiving several types of intermodal relation. Newborns can detect “arbitrary” auditory-visual relations that were presented during a period of familiarization (a particular shape paired with a particular sound), but most intermodal relations in the world are quite specific rather than arbitrary. An example is speech, which can be simultaneously heard and seen in a talking face. Adults’ phoneme perception is strongly influenced by watching faces, the so-called McGurk effect: When adults hear a syllable while looking at a face producing a different syllable, they tend to perceive the sound associated with the lip movements rather than the actual phoneme that they heard. Five-month-old infants are also susceptible to this effect.

Infants can use the duration of events to integrate information across modalities. In one study, infants were familiarized with pairs of checkerboards that flashed at the same rate but for different durations. These checkerboards were presented in silence. During test, infants were again presented with pairs of checkerboards, accompanied by sounds whose duration, onset, and offset corresponded to only one of the checkerboards. Infants older than 6 months (but not younger) exhibited a looking preference for the matching checkerboard. However, if onset and offset of the sound and flash were not synchronized (but maintained identical absolute duration), the preference disappeared, indicating that accurate responding to amodal duration information depends on synchronization of onset and offset. (*Amodal* refers to a single event that is specified by more than one modality.) Infants also may be capable of abstracting amodal rhythmic structure from auditory-visual pairings. At 5 months of age, infants can detect changes in regularly or irregularly occurring rhythmic auditory or visual sequences regardless of whether the modality of presentation is changed.

Another example of developments in intermodal perception involves matching a shape that is perceived both visually and haptically (i.e., by touch). This ability does not appear functional in very young infants, but some studies indicate that, by 4 to 5 months, infants recognize and discriminate objects cross-modally—that is, visual exposure leads to haptic recognition and vice versa.

In sum, many studies suggest that at an early age, infants are sensitive to some intermodal relations. Other findings suggest that intermodal perception for amodal pairings emerges only after 6 months or older, suggesting that at least some intermodal relations are learned through experience, or at least come on line only after particular modalities reach certain developmental levels.

HoJim Kim and Scott P. Johnson

See also Amodal Perception; Attention: Theories of; Audition; Binding Problem; Color Perception; Constancy; Cross-Modal Transfer; Eye Movements: Behavioral; Face Perception; Infant Perception: Methods of Testing; Motion Perception; Multimodal Interactions: Color–Chemical; Multimodal Interactions: Neural Basis; Multimodal Interactions: Pain–Touch; Multimodal Interactions: Spatial Perception in Touch and Vision; Multimodal Interactions: Tactile–Auditory; Multimodal Interactions: Thermal–Chemical; Multimodal Interactions: Visual–Auditory; Multimodal Interactions: Visual–Haptic; Nature and Nurture in Perception; Object Perception; Perceptual Development: Attention; Perceptual Development: Color and Contrast; Perceptual Development: Face Perception; Perceptual Development: Hearing; Perceptual Development: Imitation; Perceptual Development: Infant Music Perception; Perceptual Development: Intermodal Perception; Perceptual Development: Object Perception; Perceptual Development: Speech Perception; Perceptual Development: Taste and Olfaction; Perceptual Development: Touch and Pain; Perceptual Development: Visual Acuity; Perceptual Development: Visually Guided Reaching; Perceptual Development: Visual Object Permanence and Identity; Perceptual Organization: Vision; Speech Perception; Statistical Learning; Vision

Further Readings

Gottlieb, G., & Krasnegor, N. A. (1985). *Measurement of audition and vision in the first year of postnatal life: A methodological overview.* Norwood, NJ: Ablex.

James, W. (1890). *The principles of psychology* (Vol. I). New York: Dover.

Johnson, S. P. (2005). Building knowledge from perception in infancy. In L. Gershkoff-Stowe & D. Rakison (Eds.), *Building object categories in developmental time: 32nd Carnegie symposium on cognition* (pp. 33–62). Mahwah, NJ: Lawrence Erlbaum.

Jusczyk, P. W. (1997). *The discovery of spoken language.* Cambridge: MIT Press.

Kellman, P. J., & Arterberry, M. E. (1998). *The cradle of knowledge: Development of perception in infancy.* Cambridge: MIT Press.

Kuhl, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics, 50,* 93–107.

Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin, 126,* 281–308.

Lickliter, R., & Bahrick, L. E. (2000). The development of infant intersensory perception: Advantages of a comparative convergent-operations approach. *Psychological Bulletin, 126,* 260–280.

Saffran, J. R., Werker, J., & Werner, L. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In R. Siegler & D. Kuhn (Eds.), *Handbook of child development* (pp. 58–108). New York: Wiley.

Teller, D. Y., & Movshon, J. A. (1986). Visual development. *Vision Research, 26,* 1483–1506.

Werker, J. F., & Tees, R. C. (1999). Experiential influences on infant speech processing: Toward a new synthesis. In J. T. Spence (Ed.), J. M Darley, & D. J. Foss (Associate Eds.), *Annual review of psychology* (Vol. 50, pp. 509–535). Palo Alto, CA: Annual Reviews.

INFANT PERCEPTION: METHODS OF TESTING

Scientists once thought that infants perceived the world as a "blooming, buzzing confusion." One reason scientists now know this idea is wrong is that testing methods have been devised to ask questions about what infants see, hear, smell, and taste. This entry describes some of the most important methods of testing in the study of infant perception.

Infants and nonhuman animals pose similar challenges to a perceptual scientist: neither can follow instructions, respond to requests, or maintain interest in a task for extended periods. Accordingly, many methods of testing used to measure adult human perception cannot be used with infants. However, infants are capable of changing where they look, moving their arms and legs, and other simple actions. Moreover, their bodies and brains respond to perceptual signals in ways similar to adults. So, methods of testing infants involve measuring behaviors infants can perform or physiological responses infants reliably make. Each of these is discussed in turn.

Measuring Shifts of Gaze

Infants move their head and eyes at birth, and these shifts are readily observable by experimenters. Most of what is known about perception in infants derives from studies that have examined patterns in infants' eye and head movements, called shifts of gaze.

Robert Fantz pioneered the use of this technique. Fantz showed that both infant monkeys and humans controlled where they looked and for how long, looking more often and longer at some patterns than at others. He termed this *preferential looking,* suggesting that differences in looking direction patterns or times revealed infants' preferences for one display over another. In one demonstration, Fantz placed his participants face up in a hammock, inside of a test chamber. In the ceiling of the chamber were two holes, placed side by side. A card with a particular pattern could be placed within the left or right hole, while the other hole was indistinguishable from the ceiling background. The patterns were presented for a fixed duration several times in a row, and each time, the left or right position of the patterns was changed. A naive observer peering from outside of the chamber made judgments about where the infant looked and for how long based on the positions and patterns of a reflection from the infant's own cornea. At the end of the study, the experimenter compared the average fixation times to different types of displays. From these sorts of data, Fantz showed that infants preferred visually complex figures to simple ones and that visual acuity, based on responses to black and white striped gratings of various sizes, developed rapidly in the first few months of life.

Davida Teller made Fantz's technique even more rigorous in creating the *forced-choice preferential looking* (FPL) method. Teller instructed adult observers to make a simple left or right judgment about the direction of infant fixation and gave trial-by-trial feedback about the decisions. With this technique, it became possible to estimate psychophysical functions in infants and, thus, to determine how sensitivity changed over development. Teller invented a system for quickly and efficiently testing infants' visual acuity in the lab or clinic, called the *Teller Acuity Card* system. As a result, it is now known how infants' acuity should change across time if the visual system is developing normally and, by extension, when there is a visual disorder requiring treatment.

Fantz also demonstrated *habituation,* or a decline in the duration of looking at a familiar pattern relative to a novel one. Fantz used a method similar to his preferential looking paradigm to demonstrate this phenomenon. Instead of showing one pattern card at a time, however, Fantz first showed infants one particular pattern repeatedly until the infant no longer looked at the pattern, thus demonstrating habituation. He then presented this familiar pattern alongside a novel pattern chosen at random. Infants usually looked longer at the novel pattern. This showed that infants were quickly learning about the characteristics of the familiar pattern, but that changes in their looking times at test could provide experimenters information about which things infants perceive as different from one another.

Later investigators helped standardize the looking time habituation method, and as a result, helped make it a dominant methodology. Today, most investigators use the infant-control habituation technique that presents infants with a variable number of familiar displays before presenting both familiar and novel displays in alternating fashion in the test phase. The test phase begins when looking times have declined by an amount based on an average of the first two or three familiarization trials. Infants show novelty preferences by looking longer to test displays relative to familiar ones.

A variation on the Fantz method is the *visual paired comparison* (VPC) technique. In the VPC, pairs of stimuli are shown for fixed periods (5–10 seconds), and the proportion of looking time spent on one display versus the other is calculated. Often, the displays are identical during the initial familiarization phase, and one of the displays differs during the test phase. Gaze or head shift techniques have also been used to study auditory and olfactory perception.

All of the gaze measurement techniques mentioned thus far rely on careful observation by the experimenter coupled with film or videotape records. Recent technical advances have made it possible to record eye and gaze direction automatically with high precision using infrared sources and camera/computer systems that compute eye position from the subsequent corneal reflection patterns. These eye-tracking systems permit experimenters to examine infants' looking patterns with much greater precision. Although infants' unpredictable, erratic movements can make this method of measurement difficult to set up and calibrate, eye tracking potentially reveals details about infant perception that less precise methods cannot.

Although powerful, measures of gaze-shift behavior have shortcomings. One concerns the conditions under which infants will show novelty preferences versus preferences for the familiar. Infants, particularly younger ones, sometimes demonstrate a preference for familiar displays over novel ones or for complex stimuli. Another shortcoming relates to the fact that infants do not always look at a display that is well above their perceptual threshold; this implies that for some unknown fraction of time, infants respond based on their internal state rather than on a display's actual perceptibility. Finally, some investigators have raised questions about whether existing habituation techniques mask important individual differences.

Measuring Sucking, Kicking, Reaching, and Crawling

Gaze shifts are only one of the actions in an infant's repertoire. Experimental paradigms have been developed that tap infants' abilities to suck, kick, and, several months after birth, crawl or walk.

In non-nutritive sucking paradigms, infants are given a rubber nipple to suck; the nipple is connected to equipment that measures sucking pressure and frequency. Changes in the rate and magnitude of sucking can be used to measure perceptual discrimination. In one example, experimenters measured infants' sucking pressure to speech sounds that were presented repeatedly. Then, experimenters measured how the pressure changed when a novel sound was presented. In other experiments, scientists used visual feedback from a slide projector connected to the apparatus to reinforce infant sucking responses. Infants in the study rapidly learned to control the focus of the slide projector by sucking faster or more slowly to bring the image into clear view. This suggested that infants are motivated to perceive the external world and will act to achieve perceptual goals.

Experimenters have also measured infant kicking responses to study perception. Infants rapidly learn to vary the frequency of kicking or leg movements when those actions evoke other events, such as when the arm or leg is connected to an overhead baby mobile by means of a ribbon. Scientists have shown that kicking will be suppressed when infants perceive a change in the mobile's visual characteristics or those of the crib or room where the mobile is presented. Moreover, by varying the delay between training and test and measuring infants' movements, researchers have shown that infants can remember specific mobiles for hours, days, or even weeks.

Motor responses have also been used to study the development of depth perception. Based on the assumption that infants will reach toward objects that appear closer than those that appear far away, scientists have shown that sensitivity to some depth cues emerges between 5 and 7 months. Infants' crawling behavior has also been used to study depth perception. Eleanor Gibson and Richard Walk created a simulated "visual cliff" over which babies were encouraged to crawl. The visual cliff apparatus consisted of a large box with two sides, one shallow and one deep. An identical checkerboard texture covered both sides, and a sturdy glass plank was placed on top to create a solid surface for crawling, but with the visual impression of a steep cliff on one side. Infants were placed on the shallow side and called by their mothers to crawl to the deep side. When infants reached the edge of the cliff, most stopped crawling, even when their mothers repeatedly called to them from the other side. The study demonstrated that most human infants perceive depth by the time they begin to crawl. More recent studies have used psychophysical techniques that vary the slope of the drop-off in an effort to determine how infants perceive which slopes are too steep and which are not. Intriguingly, information learned about slopes from crawling does not immediately transfer when infants start to walk.

Measuring Physiological Activity

Direct measurements of physiological activity provide powerful insights into infants' perception that can complement behavioral techniques. *Electrocardiography* (ECG) is a technique to measure heart rate. Heart rate changes in response to both physical and psychological states. Changes in heart rate have been used to measure infants' speech sound discrimination and even fear responses on the visual cliff. *Electroencephalography* (EEG) is used to measure electrical activity of the brain via an array of electrodes placed on the scalp. Perceptual studies incorporating EEG usually focus on evoked cortical responses to the presentation of a stimulus. The most common method to study evoked activity through EEG is by computing the *evoked response potential,* or ERP. In perceptual ERP studies, a stimulus is presented repeatedly. Numerous stimulus-locked trials are recorded and averaged together to yield the ERP pattern for a particular stimulus type. The amplitude and timing of the evoked response is computed in reference to the presentation of the stimulus. There is also a large literature on *visual evoked potentials* (VEPs) in infants. The *steady-state visual evoked potential* (SSVEP) method is one such technique. In SSVEP, a stimulus changes (either on/off or reversing) at a fixed temporal frequency or rhythm, and the evoked EEG responses are analyzed by focusing on the specific frequencies in the stimulus. SSVEPs have been used to study the development of many different perceptual qualities, from visual acuity to complex motion processing. Evoked potential techniques can be used to study other sensory systems. The *brain-stem auditory evoked potential* (BAEP) is a well-established indicator of central nervous system auditory function that is used frequently in clinical settings to determine whether an infant's auditory system is functioning normally.

Implications

The methods available for testing hypotheses in a rigorous experimental setting constrain progress in many areas of science. Pioneers such as Fantz and Teller have opened up new vistas of research in infant perception as a result of the careful development and promotion of clever experimental techniques that allow scientists to ask infants what they can see, hear, taste and smell. As a result, researchers know that infants occupy a perceptual world that differs from adults' in certain respects but is nonetheless full, rich, textured, engaging, and above all, informative.

Rick O. Gilmore and Jeremy D. Fesi

See also Evoked Potential: Audition; Evoked Potential: Vision; Infant Perception; Psychophysics: Detection; Visual Acuity

Further Readings

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171,* 303–306.

Fagan, J. F. (1990). The paired-comparison paradigm and infant intelligence. *Annals of the New York Academy of Sciences, 608,* 337–357.

Fantz, R. L. (1964). Visual experience in infants: decreased attention to familiar patterns relative to novel ones. *Science, 146,* 668–670.

Gibson, E. J., & Walk, R. D. (1960). The visual cliff. *Scientific American, 202,* 67–71.

Siqueland, E. R., & DeLucia, C. A., (1969). Visual reinforcement of non-nutritive sucking in human infants. *Science, 165*(3898), 1144–1146.

Teller, D. (1979). The forced-choice preferential looking procedure: A psychophysical technique for use with human infants. *Infant Behavior and Development, 2,* 135–153.

INFORMATION THEORY

Some theories excite the popular imagination so strongly that they acquire broad cultural influence. Ideas and terminology from the technical core reach significantly beyond the original domain of application. *Information theory* is one such body of ideas, canonized by the name we give to our times ("the age of information"). The initial developments of information theory, notably at Bell Laboratories, were undertaken by engineers and mathematicians such as Harry Nyquist and David Slepian trying to rationalize the design of communication systems and drew from concepts from statistical physics and signal processing. Claude Shannon, in a series of papers beginning with "A Mathematical Theory of Communication," gave the field definitive shape and used similar ideas to develop a theory of cryptographic systems. Important contributions were also made by Norbert Wiener and Andrei Kolmogorov. The resulting framework remains one of the most conceptually well-characterized engineering theories.

Note that the phrase "information theory" can be used with a broad or a narrow scope, depending on taste: It is used in this entry in a relatively narrow way, to refer to the "mathematical theory of communication," as Shannon did. Concepts from information theory and signal processing enter into the study of perception in two ways. The first is to study some aspect of perception from the point of view of a communication engineer: How good are our visual, auditory, gustatory, olfactory, and other "sensors"? Do they pick up signals with high fidelity, and how rapidly can they sense changes in the environment? The second, related usage is in statistical measures for analyzing recordings of behavior or neural activity.

The original concern of information theory was to understand the communication of messages over noisy channels. Consider a binary channel in which only one message can be sent every time slot, with two possible values ("yes," "no"). Now suppose that the channel is noisy, so that sometimes a "yes" becomes a "no" during transit through the channel, and vice versa, with probability p. If this happens relatively rarely ($p < 1/2$), one could still send a reliable "yes" message through the channel, by sending a long block of "yes"s. Some of these will become "no" because of the noise process, but by taking the majority vote, one should be able to recover the original message. For a finite block length, there is always a possibility that a chance fluctuation will lead to the wrong majority vote—but this possibility of error should intuitively decrease as one takes a longer and longer block. Mathematically, the law of large numbers can be used to show that as long as the probability of changing a "yes" message to a "no" message is less than one half, the probability of obtaining the wrong message by taking majority vote will go to zero as the block size becomes infinitely large. This corresponds to sending an unambiguous, error-free message over a noisy channel.

There is a catch with this strategy, however: Because infinitely long blocks are needed to make a single "yes" message truly noise free, it would take infinitely long to send a single noise-free binary message. This does not work because the rate of error-free communication (which is proportional to the inverse of the time taken to send a single error-free binary message) then drops to zero. Shannon's insight was to realize that by appropriately introducing redundancy into a time varying signal, error free communication was possible at a *finite* rate per unit time, called the information *capacity* of the channel.

This capacity can be shown to be the difference between the logarithm of the number of possible output messages (which is also called the *entropy* of the message set), and the logarithm of the number of possible output messages if the input message were held fixed. If the base of the logarithm is 2, then the "unit" of measurement is a "bit." For the binary channel described, the number of possible messages per unit time is 2, corresponding to a message entropy of $\log(2) = 1$ bit per unit time, and the channel capacity can be shown to be $C(p) = \log(2) + p\log(p) + (1 - p)\log(1 - p)$. If $p = 1/2$, then this capacity falls to 0. This makes sense because at that point, a "yes" is converted to "no" about half the time, and the message is completely garbled. However, for $p < 1/2$, the theory effectively says that once in every $1/C(p)$ time units, a noise-free "yes" or "no" message can be sent. This surprising result is achieved through the usage of *error-correcting codes,* the design of which forms an important component of information theory.

Real channels operate in continuous time, with continuous valued signals. These signals cannot vary arbitrarily rapidly and have a finite *bandwidth* (the effective number of samples per second). This is a concept from signal processing that is important both for understanding perception and for designing corresponding technologies: familiar examples include digital audio (CD quality recordings use a bandwidth of 44 kilohertz [kHz], consistent with the bandwidth limitations of human audition, which is in the 10–20 kHz range) and cinema (film traditionally used 24 frames per second, consistent with the bandwidth of human vision). Because of limited bandwidth, we cannot perceive fluctuations in the visual scene that are sufficiently rapid. One can also ask how many *distinct levels* may be discriminated within an individual time bin (recall the two-levels per time bin in the example of a binary channel). This may be estimated via psychophysical experimentation in which subjects are asked to discriminate between levels of intensity, for example. Once these two measurements that characterize the perceptual channel are at hand, it is tempting to define the "channel capacity" using the formulation described, and indeed this has been done. There are some conceptual difficulties, however: It is questionable whether there is a well-defined "sender" and an encoder in the external world sending messages to a receiver inside the nervous system—such a sender-receiver pair would be nominally required to define channel capacity. A different and perhaps better grounded usage of these concepts is to the communication between two organisms, where sender and receiver are well defined. Signal multiplexing schemes analogous to those used in cellular telephony may be found in frogs and electric fish, and information theoretic ideas have been applied fruitfully to the speech channel in humans.

Single-user communication theory, as described, is a mature subject area. The same cannot be said of multi-user communication theory or network information theory, where there are multiple senders and receivers with varying degrees of interference and cooperation. This area of contemporary research promises to be of significant utility in understanding biological function. Finally, the reader will be well advised to remember Shannon's cautionary note about information theory having become something of a "scientific bandwagon": if it was already true in Shannon's days, it is certainly more so today.

Partha Pratim Mitra

See also Psychophysics: Detection; Signal Detection Theory and Procedures

Further Readings

Shannon, C. E. (1948, October). A mathematical theory of communication. *Bell Systems Technical Journal, 27,* 379–423, 625–656. Retrieved from http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf

Shannon, C. E., & Weaver, W. (1998). *The mathematical theory of communication.* Urbana: University of Illinois Press.

Verdú, S., McLaughlin, S. W., & IEEE Information Theory Society. (2000). *Information theory: 50 years of discovery.* New York: IEEE Press.

INTENTIONALITY AND PERCEPTION

As you drive your car, the focus of your perceptions shifts rapidly between the other cars on the road, the song playing on the radio, and the conversation you are having on your cell phone. In each case, your perceptions are "about" something. Perceptual scientists and philosophers have long referred to this "aboutness" quality of perception as its *intentionality*.

Early in psychology's history (i.e., the late 1800s), perception and intentionality were conceptualized in terms of their relationship to consciousness. Thus, perceptions were treated as conscious experiences that were inherently intentional because they were directed toward (i.e., "about") some object. In the early 20th century, psychology moved away from studying consciousness and began to focus on observable behavior. As a result, perception and intentionality underwent somewhat of a split as the notion of "intentionality" came to be associated with the goal-directed behavior studied by behaviorists, and perception came to be treated as the conscious responses to stimuli studied by psychophysicists.

With the advent of cognitive science and information-processing theory (IPT) in the 1940s, researchers began studying our ability to attend to some stimuli and filter out others, as we do, for example, when trying to carry on a cell phone conversation in a loud room. Instead of referring to this ability to direct our perceptions toward specific stimuli as an act of "intention," however, researchers referred to it as an act of "attention." As a result, intentionality and perception were completely split, and cognitive scientists came to describe behavior as intentional (e.g., "He intended to press the button") but not perception (i.e., one would say, "He attended to the stimulus," rather than, "He intended to see the stimulus").

Within the IPT framework, many researchers worked to discover the internal processes responsible for transforming raw sensory input into detailed perceptual representations that could be used to direct behavior. This approach was strongly influenced by the advent of cognitive neuroscience and the discovery that neurons in the initial visual processing areas of the brain (i.e., V1 in the occipital cortex) seem to function as *edge detectors*. That is, certain V1 neurons responded most actively to stimuli that had the shape of an edge. Later research revealed neurons in higher visual areas (e.g., V2, V4, and V5) that had even more specific response properties (e.g., responding maximally to motion in a particular direction). Such discoveries strengthened IPT's treatment of perception as the internal processing of representations to be used for generating behavior. This, in turn, strengthened the association of intentionality with behavior, and attention with perception.

During this time, a group of researchers known as *ecological psychologists* rejected the idea that the purpose of perception is to construct internal representations of stimulus properties such as size, shape, and color. Instead, ecological psychologists proposed perception is about detecting the behavioral possibilities a stimulus affords. Thus, rather than perceiving a stick as an object of a certain length and color, ecological psychologists claim we perceive it in terms of its affordances—what we can *do* with it. For example, research indicates people perceive whether they are able to walk up a step of a particular steepness, and the step height that constitutes the boundary between "yes" and "no" is actually a constant relationship between the height of the step and the perceiver's leg-length. From the ecological perspective, this means the perception of the steps is in terms of the relationship between "behavioral" and "stimulus" properties, rather than just "stimulus" properties alone. This approach begins to reunite perception with intentionality because it conceptualizes perception as being "about" behavior (as opposed to being "about" stimulus properties that are then used to guide behavior).

Recently, hybrids of the ecological and IPT positions have emerged that might ultimately integrate the two. In cognitive psychology, for example, researchers have proposed the *theory of event coding* (TEC), which asserts the following: (a) actions are planned for the external events they are to produce, and (b) action planning and perception share neural resources. This position renders perception inherently intentional because it implies that as soon as one plans to produce an external event such as crack a peanut or pick up a cup, areas of the brain involved in planning the event are also used to perceive the event. This assertion has recently been supported by neurophysiological findings that indicate that neurons in the premotor areas of the brain (known as *canonical neurons*) become active when a monkey simply sees an object that affords actions such as grasping and tearing, as well as when the monkey executes such actions toward the object. Further, other neurons in this same premotor area, known as *mirror neurons,* become active when the monkey sees another actually interact with the object (or when it interacts with the object itself).

Collectively, canonical and mirror neurons are consistent with the idea that perception is inherently intentional because perception and planning share overlapping neural resources. Premotor areas of the brain that are active as we plan to produce an event are also active when we perceive such events. In short, planning and perceiving are, to some extent, the same thing. Although canonical neurons seem to underlie the perception of what we *can* do (what the ecological psychologists refer to as *affordances*), mirror neurons seem to underlie the perception of what others plan to do. That is, by perceiving the actions of others in terms of the plans we would use to produce the actions ourselves, we are, in effect, perceiving their intentions. Many researchers are now proposing that the mirror system constitutes the means by which we have access to other minds.

Given the emergence of these hybrid approaches, it seems perception and intentionality may have come full circle. In the late 1800s, perception and intentionality were considered part of the same phenomena; namely, directed, conscious experience. Now, with the ecological focus on affordances, the emergence of TEC, the discovery of canonical and mirror neurons, and the realization that perception can be in terms of behavioral possibilities and plans, perception and intentionality seem to once again be part of the same phenomenon. However, it is still not clear to what extent perception is always in terms of possibilities and plans. Research indicates, for example, that prefrontal areas of the brain that become active when one thinks about one's own thoughts also become active when one thinks about the thoughts of another, but not when one thinks about the processes of a computer. It might be, therefore, that we can *choose* whether to perceive a phenomenon either in terms of behavioral possibilities and plans or objective stimulus properties. If this proves true, however, the ability to choose between perceiving something objectively (i.e., in terms of stimulus properties) versus subjectively (i.e., in terms of plans) will serve as evidence that perception is more inherently intentional than we ever imagined.

J. Scott Jordan

See also Gestalt Approach; Phenomenology (Philosophy); Philosophical Approaches; Theoretical Approaches

Further Readings

Gallagher, H. C., Jack, A., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance. *Neuroimage, 16,* 814–821.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24,* 849–937.

Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *Journal of Neurophysiology, 28,* 229–289.

Jordan, J. S. (2008). Wild-agency: Nested intentionalities in neuroscience and archeology. *Philosophical Transactions of the Royal Society B (Biological Sciences), 363,* 1981–1991.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (2002). From mirror neurons to imitation: Facts and speculations. In A. Meltzoff & W. Prinz (Eds.), *The imitative mind* (pp. 247–266). New York: Oxford University Press.

Warren, W. H. (1984). Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 683–703.

INVERTED SPECTRUM

Suppose persons A and B have different color sensations while gazing at a ripe red tomato: A perceives "redness," but B has the sensation that A has when A is looking at a purple grape. A and B both can visually detect ripe tomatoes and both learned to use language the same way; only their subjective sensations differ. If A and B have systematically swapped sensations (red-purple, orange-blue, etc.), then they are said to have an inverted qualia spectrum relative to each other. Thus, the *inverted spectrum* (IS) possibility is that, given the same wavelength light stimulus, two people might systematically have different color sensations (qualia) yet behave the same.

The IS received its classic statement in John Locke's 1690 *An Essay Concerning Human Understanding*. Locke considers inversion between subjects; later authors (perhaps Ludwig Wittgenstein first) have considered inversion at different times within a single subject. The IS continues to raise serious problems for a range of theories of language, mind, and consciousness, as discussed in this entry.

Significance for Philosophy and Psychology

Behaviorism claimed a solution to the problem of knowing other minds—mental states are entirely evident in behavior. On one form, to be in mental state M is just to be disposed to exhibit behavior R in response to stimulus S. Thus, to sense red is to respond to long light wavelengths by saying "that's red" or "that's a ripe tomato." But the IS raises an obvious problem: A purple sensation is on the face of it a different mental state than a red sensation is, yet an inverted spectrum would not affect overt behavior. Thus, behaviorism, which sought to reduce all mental states to response to stimuli, fails to account for qualitative mental states. Different states could cause the same behavior; the mental does not reduce entirely to stimulus-behavior links.

Functionalism and computationalism were successors to behaviorism developed starting in the 1960s. A functional account characterizes a mental state by its causal role: its inputs and outputs to other mental processes, in addition to the observable stimulus and response to which behaviorism was limited. For example, pain is that state that is typically caused by injury and that in turn might cause a mental planning module to avoid pressure on an injured limb. For a functionalist, a pain state is any state that has this causal role. But again, the inverted spectrum creates a problem: differing qualia states (red, blue) could have the same functional role in different individuals. Functionalism seeks to reduce each mental state to its causal role in a cognitive system, but the IS seems to show that different sensations could play the same causal role.

Indeed, perhaps normal behavior could be produced with no subjective consciousness at all. The IS suggests that the character of sensations doesn't matter and is undetectable; hence, an even more radical possibility than the IS is that some people have no subjective experiences at all—they behave the same in response to stimuli as normal conscious subjects but experience no qualia at all (while avowing that they do). This extension of the IS is the *absent qualia* or *zombie* possibility raised by Ned Block in the 1970s and widely discussed since.

The inverted spectrum also raises problems for theories of language and linguistic meaning, including Locke's own. Locke held that words stand for ideas in the minds of language users. For example, *horse* stands for the idea of horse, which Locke seems to have understood to be a mental image. *Red* stands for the sensation of red. But if ideas and sensations can vary undetectably between subjects, as the IS seems to show, then meanings can vary undetectably between persons. If the point of language is to communicate meaning, there is no way to know if communication is successful if meanings are intrinsically private subjective mental states. Wittgenstein argues in the "private language argument" that this shows that a mentalistic semantics held by Locke and later empiricists such as Bertrand Russell is incoherent. Despite this, prominent contemporary theories of meaning (Jerry Fodor, Steven Pinker) hold that linguistic meaning derives from mental states. These contemporary language theorists appeal to innateness claims unavailable to blank slate empiricists such as Locke—but verifying the claim that all humans have the same innate mental states remains a problem.

Empirical Issues

It has been argued that there are natural partial IS phenomena. Martine Nida-Rumelin argued that known genes cause our light-sensitive pigments to be produced in the "wrong" cone; hence, there could be combinations of these genes that cause the pigments to be "switched" with no manifest behavioral effects. However, this suggestion presupposes that experienced qualia are determined by low-level physiology, rather than by what is done with the information by higher more integrative and holistic processes. Others have argued that known asymmetries in the color space imply that a simple inverted spectrum is psychologically impossible.

David Cole argued that an inverted spectrum might be experimentally produced using electronic color inverting goggles. Past experiments by George Stratton and Ivo Kohler with simple visual field inverting goggles suggest that after prolonged wear, the subjective visual field reverts to right side up despite inverting the image on the retina. Experiments by Celeste McCollough and others in the mid-1960s with partially tinted contact lenses that alter the color of parts of the visual field also suggest that acclimation occurs and the subjective states revert to normal. Thus, Cole suggests that a possible outcome of color spectrum inverting goggles would be that with behavioral adaptation would come qualia reversion. If qualia revert with adaptation, the acquired inverted spectrum possibility would not be a threat to functionalism or computationalism.

In 1990, Block devised "Inverted Earth," a place where all external colors are inverted with respect to Earth (the sky is red, etc.) to allow inverts to not notice their inversion. But perhaps connection to the world and color proximity are not the only dimensions in which colors vary. Cole subsequently argued that colors vary in emotional associations (e.g., red is alarming, blue soothing), and suggests that affect may play an important role in the subjective character of colors. To preserve functional role after inversion, the emotional color of color would need to be altered as well, but the connection may be innate, immutable, and partially constitutive of the experience. Further, Cole considers other inversions, including inversion of auditory pitch and perceived loudness and a variant on synesthesia: the radical scenario of vision producing qualia from different sense modalities in differing subjects (tranesthesia, e.g., B has auditory sensations where A has visual; with no differences in overt behavior), and argues that it is implausible that if adaptation is possible with these inversions, qualia would remain unaffected.

Thus, the inverted spectrum and related thought experiments are argued to have a range of theoretical and empirical implications and will likely be the subject of lively debate into the future.

David Cole

See also Mary the Color Scientist; Philosophical Approaches; Private Nature of Perceptual Experience; Qualia

Further Readings

Byrne, A. (2006). Inverted qualia. Retrieved from http://plato.stanford.edu/entries/qualia-inverted

Cole, D. (2002). Inverted spectrum arguments. Retrieved from http://www.d.umn.edu/~dcole/inverted_spectrum.htm

Locke, J. (2008). *An essay concerning human understanding.* New York: Oxford University Press. (Original published in 1690)

Putnam, H. (1981). *Reason, truth, and history.* New York: Cambridge University Press.

Stratton, G. (1897). Vision without inversion of the retinal image. *Psychological Review, 4,* 341–360, 463–481.

ITCH, TICKLE, AND TINGLE

Itch, tickle, and tingle are commonplace sensory experiences that form part of the larger constellation of skin senses that also includes touch, temperature, and pain. Itch and tickle are sometimes linked perceptually, but under most circumstances these two distinct sensations are evoked by markedly different stimuli: chemicals for itch, and touch stimulation for tickle. This entry describes these sensations.

Itch

Itch is defined as an unpleasant sensation associated with the desire to scratch. Itch is commonly evoked by insect bites, allergic reactions, or contact with certain plants such as poison ivy or cowhage, a bean plant with seedpods that are covered with spicules that cause itch when they contact skin. Allergic reactions such as hives are caused by the release of histamine in the skin, which elicits a nearly pure sensation of itch and associated reddening of the skin (flare). Histamine excites a particular class of thin (unmyelinated), touch-insensitive nerve endings that transmit impulses to the spinal cord and, from there, to higher centers in the brain to result in itch sensation. Itch caused by histamine release is usually relieved by antihistamines. Interestingly, itch from cowhage does not involve histamine but, rather, a different chemical called mucunain. Mucunain is contained within the cowhage spicules that cause itch without a flare reaction when they contact the skin. Mucunain is a protease that interacts with protease-activated receptors (PARs) that reside on sensory nerve endings in the skin. It is currently under investigation if mucunain activates a separate class of touch-sensitive nerve endings in the skin that transmit itch independently of histamine. The existence of a histamine-independent itch pathway has great significance for patients who suffer from chronic itch caused by liver and kidney disease, AIDS, and many other diseases and skin conditions in which the itch is not relieved by antihistamines. A better understanding of this type of itch will hopefully lead to new and improved treatments for chronic itching.

Itch is often considered to be a low level of pain sensation. However, several observations suggest that itch is distinct from the sensation of pain. Itch warns the organism of insects or plants contacting the skin surface, which can then be protectively removed by scratching or rubbing. In contrast, a painful stimulus usually evokes protective withdrawal of the stimulated skin area away from the stimulus, rather than scratching, which would worsen the pain. Scratching is itself a mild form of pain that relieves itch. Morphine, which reduces pain, frequently causes itching. Experimental studies have shown that electrical stimulation at certain discrete sites on the skin can evoke a sensation of itch; increasing the frequency of stimulation evokes a stronger itch sensation that never becomes painful. Despite these distinctions between itch and pain, it is still not certain if they are conveyed by their own separate pathways. Both painful and itchy sensations activate largely overlapping brain regions including the somatosensory and supplementary motor cortex, insula, and anterior cingulate gyrus. Furthermore, although pain, itch, and temperature are all conveyed primarily by the spinothalamic tract, current evidence suggests that itchy and painful stimuli excite a common set of neurons rather than distinct subpopulations of itch- and pain-specific neurons. Increased research interest in this question could answer whether itch and pain are conveyed by separate or shared pathways.

Tickle

In contrast to itch, tickle sensation is elicited by touch. Some time ago, psychologists distinguished between two types of tickle. The first type, called knismesis or light tickle, is the well-known sensation elicited by slow movement of a light object, such as a feather, across the skin. This sensation can be pleasant, but it can also be unbearable to the extent that the tickled limb is moved away, similar to withdrawing from a painful stimulus. An example is the tickle sensation evoked by lightly stroking the upper lip or the sole of the foot. The second type of tickle, called gargalesis or heavy tickle, is the familiar sensation frequently associated with laughter that is evoked usually by playful (but sometimes cruel) and more forceful stimulation of the body. This type of tickle has been contemplated by Aristotle and Charles Darwin, who noted that one cannot tickle oneself and that the emotional valence of heavy tickle depends on the social setting, being pleasant when the tickler is a familiar friend or family member but threatening and even torturous if the tickler is a stranger.

The inability to tickle oneself has been investigated using modern brain-imaging techniques. When subjects moved one arm to tickle their own hand on the opposite side, there was significantly less activation of relevant areas of the brain (secondary somatosensory and anterior cingulate cortex) compared with when an experimenter tickled their hands. The anterior cerebellum, a brain structure associated with coordinated movement, is also active when subjects tickle their own hands. Interestingly, activity in the cerebellum was greater when the subject moved his arm without tickling his other hand, compared with when he moved his arm and actually tickled the other hand. This suggests that the cerebellum normally predicts the sensory consequences of a self-generated movement, and that this prediction can actually cancel the brain's response that is associated with the sensation normally elicited by the movement. In other words, a self-generated tickling movement cancels the tickle sensation, explaining why we cannot tickle ourselves.

Is tickle a unique quality of sensation, as pain and itch are thought to be? If so, tickle receptors should exist in the skin. Decades ago, neurophysiologists discovered thin unmyelinated nerve fibers that responded to light stroking of the skin and suggested that these may be "flea" or tickle receptors. Such receptors have been found more recently in human skin. However, the role of such fibers in tickle has been challenged by a unique patient with a verified loss of large-diameter myelinated nerve fibers following an illness. This patient could still detect temperature, pain, and itch, which are conveyed by the intact unmyelinated fibers. However, she was insensitive to tickle and virtually insensitive to touch because of the loss of myelinated fibers. When carefully tested with brushing of the skin, she reported the sensation to be pleasant. This interesting case suggests that the unmyelinated touch-sensitive fibers do not convey tickle, but instead some of them may convey pleasant touch sensation. It also suggests that light tickle is conveyed by larger myelinated nerve fibers. In this regard, it is interesting to consider studies involving electrical activation of single functionally identified nerve fibers (intraneural microstimulation) in conscious humans. Activation of a nerve fiber along its path from the skin receptor to the brain evokes the same sensation that would result if the receptor had actually been physically touched. Subjects have occasionally reported tickle sensation in response to stimulation of individual nerve fibers connected to light touch (rapidly adapting) or vibration (Pacinian) receptors. Based on these findings, it is possible that light tickle is conveyed by specific, but as yet unidentified, receptors having large myelinated nerve fibers. Alternatively, light tickle may be a complex sensation that depends on a particular spatiotemporal pattern of activity in touch receptors.

Tickle has been described as a moving itch, which by definition is unpleasant. However, lightly stroking the skin can evoke pleasant or unpleasant sensations, probably dependent on one's mood as well as other poorly understood factors. Moreover, itch can be elicited by lightly touching skin in the area surrounding an itchy site such as a mosquito bite, a phenomenon called allokinesis. Itch can also be evoked by lightly touching skin in the absence of any surrounding itch. The relationships among pleasant touch, tickle, and mechanically evoked itch are somewhat blurred and not well understood.

Tingle

To further complicate matters, *tingle* and *prickle* are also commonly used to describe related but distinct skin sensations. Tingle has been best studied in relation to drinking carbonated beverages. Most carbonated drinks contain carbon dioxide bubbles that effervesce when the can or bottle is opened, producing a fizzy sensation. Although the bursting bubbles may elicit a mild touch sensation, there is also a strong tingle sensation that persists for several seconds after the drink has been swallowed or spat out of the mouth. This sensation cannot be attributed to bursting bubbles since they are no longer present. Tingle sensation is reduced or eliminated by drugs that interfere with the enzyme (carbonic anhydrase) that converts carbon dioxide gas to carbonic acid. Indeed, mountaineers taking the carbonic anhydrase inhibitor acetazolamide for altitude sickness complain that carbonated drinks no longer taste fizzy. Carbonic acid excites sensory nerve endings in the mouth, including pain receptors (nociceptors). Indeed, prolonged contact of the tongue with a carbonated beverage can result in a burning irritation or pain. Nevertheless, many humans are able to overcome these aversive properties to develop a preference for carbonated drinks. Similarly, prickle is an unpleasant "pins and needles" sensation often experienced when wearing clothing made of a coarse fabric such as wool. Prickly fabrics activate nerve endings in the skin that signal pain. Thus, like tingle, a major component of prickle sensation involves the activation of pain pathways.

E. Carstens

See also Cutaneous Perception; Cutaneous Perception: Physiology; Perceptual Development: Touch and Pain; Temperature Perception

Further Readings

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience, 1*(7), 635–640.

Garnsworthy, R. K., Gully, R. L., Kenins, P., Mayfield, R. J., & Westerman, R. A. (1988). Identification of the physical stimulus and the neural basis of fabric-evoked prickle. *Journal of Neurophysiology, 159*(4), 1083–1097.

Harris, C. R. (1999). The mystery of ticklish laughter. *American Scientist, 87*(4), 344–351.

Ikoma, A., Steinhoff, M., Ständer, S., Yosipovitch, G., & Schmelz, M. (2006). The neurobiology of itch. *Nature reviews. Neuroscience, 7*(7), 535–547.

Johanek, L. M., Meyer, R. A., Hartke, T., Hobelmann, J. G., Maine, D. N., LaMotte, R. H., & Ringkamp, M. (2007). Psychophysical and physiological evidence for parallel afferent pathways mediating the sensation of itch. *Journal of Neuroscience, 27*(28), 7490–7497.

Olausson, H., Lamarre, Y., Backlund, H., Morin, C., Wallin, B. G., Starck, G., et al. (2002). Unmyelinated tactile afferents signal touch and project to insular cortex. *Nature Neuroscience, 5*(9), 900–904.

Schmelz, M., Schmidt, R., Bickel, A., Handwerker, H. O., & Torebjörk, H. E. (1997). Specific C-receptors for itch in human skin. *Journal of Neuroscience, 17*(20), 8003–8008.

K

KINESTHESIA

When a pitcher throws a fastball to a batter, the velocity of the ball can be as high as 100 miles per hour (mph). A few milliseconds difference in the timing of ball release can make the difference between a strike and a ball. The control of the ongoing configuration of the arm and of the fingers releasing the ball thus has to be precise. Such precision is actually characteristic of most of our hand movements. For example, when we run our hands over a stationary surface, we will perceive the surface to be stationary; by contrast, if it is moving, we will perceive it to be moving and sense the direction of its motion. The ability to make such distinctions means that we are able to discriminate the stimulation of the skin of our hand that is caused by our own arm motion from that resulting from motion of the touched surface.

Kinesthesia refers to the ability to sense limb motion and position. This information is an important component in the control of voluntary motion and in skilled action. A pitcher deprived of kinesthesia would not be able to throw a baseball accurately. Similarly, we would not be able to distinguish without looking whether the surface we ran our hand over was moving or not. When our limbs are moved passively as by a neurologist during a clinical evaluation, we also have an awareness of their displacement. The kinesthetic sense is thus independent of whether a limb movement is made voluntarily or not. In the case of a voluntary movement, however, we also have a sense of effort associated with making the movement, and if we are tired, the same movement seems to require more "effort." Kinesthesia is closely related to proprioception and their underlying nervous system mechanisms are closely related as well. Kinesthesia encompasses the perception of active and passive limb motion, the sense of voluntary effort associated with a movement, and the sense of muscular fatigue. Proprioception emphasizes the ongoing static and dynamic configuration of the whole body and its orientation to gravity. This entry discusses receptors contributing to kinesthesia, phantom limbs, and the sense of effort and of fatigue.

Receptors Contributing to Kinesthesia

A wide variety of sensory information from skin, joints, and muscles contributes to our sense of limb position and movement. For many years, it was thought that "joint receptors" gave rise to the sense of limb position, acting almost like angular potentiometers with each angular position being associated with a unique neural signal. Several types of sensory endings are located in the joint capsule itself, in the ligaments of the joint, and in the fibrous periosteum (the dense connective tissue that covers bones) near the attachment of ligaments. Each of these regions has a specific and different type of specialized receptor ending whose signals are conveyed in nerves subserving the joint (known as articular nerves).

Initial studies in the 1950s recorded activity from afferent fibers in the articular nerves and related their firing properties to the position and change of position of a joint. Typically, these fibers showed maximum activity over a small range of joint angles. Some would maintain near constant discharge for prolonged periods even if the joint were not moved, suggesting that they could signal the static position of the limb. During movement of the joint, these fibers would show increased discharge for displacements in the direction of the sensitive region and decreased discharge for the opposite direction. Overall, the pattern of results from multiple studies seemed to give converging evidence that receptors associated with the joint could provide a comprehensive representation of joint angle with subpopulations of receptors coding distinct subregions of a joint's range of motion.

This simple picture became more complicated, however, when physiologists began to record from individual nerve fibers in the dorsal roots of the spinal cord, where sensory input enters the central nervous system, and to identify specific receptor types associated with the joint and periosteum. It came as a genuine surprise that most receptors were maximally active only at extreme joint angles, which could include both flexion and extension, and that only some were sensitive in the broad intermediate zone of joint motion. Moreover, relatively few in the intermediate range of limb motion continued to fire when the limb was held in a constant position. By the late 1960s, it became clear that receptors specifically associated with the joint capsule and ligamentous attachments were unlikely alone to account for the ability to detect accurately joint position and changes in joint angle.

A solution to the riddle of position sense accuracy emerged in the 1970s when research implicated receptors within the muscles and their motor innervation as being of key importance. Surprising as it may seem—because we usually think of muscles as being contractile tissue subserving our ability to move—muscle has many sensory elements. Golgi tendon organs lie in the tendons of muscles near where they attach to bones and effectively serve as exquisite tension or force sensors with their output being related to the level of tension at the tendon. Small, unmyelinated fibers, "C fibers," also lie in muscle and subserve temperature and pain sensation. The body of the muscle contains two distinct types of muscle fibers, the striated muscle fibers or extrafusal muscle fibers that do the actual force generation of muscle contraction, and the extrastriate or intrafusal muscle fibers. The intrafusal muscle fibers are quite complicated structures. They are also commonly known as muscle spindle fibers. In their central or equatorial region, they have two types of sensory endings, referred to as spindle primary and spindle secondary receptors. Spindle primaries have large diameter axons that conduct signals rapidly to the central nervous system via the dorsal roots of the spinal cord. Investigators have found that primaries respond both to static length of the intrafusal fiber and, especially, to change in length or velocity. By contrast, the spindle secondaries have smaller diameters and slower conducting axons, and their discharge rate seems related to intrafusal muscle fiber length.

A century ago, Charles Sherrington noted that there were nearly 10 times as many sensory fibers from muscle spindles as there were from the receptors associated with joints. He thought this raised the likelihood that muscle spindles were important contributors to kinesthesia. This possibility was long resisted because intrafusal muscle fibers have an independent innervation from that of the extrafusal muscle fibers. The latter are innervated by the large alpha motoneurons of the spinal cord. By contrast, the intrafusal muscle fibers are innervated by the small gamma motoneurons of the spinal cord that independently adjust the tension across the equatorial region of the intrafusal muscle fibers that lie in parallel with the extrafusal muscle fibers. If only alpha motoneurons were active, a muscle would contract and the intrafusal muscle fibers would be unloaded so that there would be no tension across their force-sensitive equatorial regions. Figure 1 illustrates the control of the intrafusal and extrafusal muscle fibers and the sensory afferents from the intrafusal primary and secondary endings.

It was long thought that muscle spindle receptors could not contribute to the sense of limb position and motion because their discharge would depend on the levels of alpha and gamma motoneuronal activity and the orientation of the limb relative to gravity (which affects the relative force necessary to move and position it). As a consequence, there cannot be a simple one-to-one relationship between alpha and gamma motoneuronal activation and activity of spindle primary

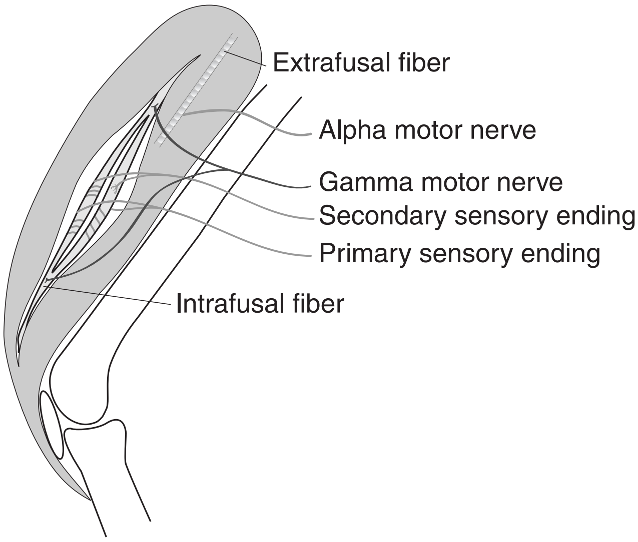


Figure 1 Arrangement of the Extrafusal and Intrafusal (Muscle Spindle) Fibers in a Striated Muscle

Notes: The gamma motor nerve controls the sensitivity of the spindle primary and secondary receptors. The alpha motor nerve controls the contraction of the entire muscle. The spindle primary axons have an excitatory influence on the alpha motoneurons controlling the muscle in which they lie.

and secondary receptors and joint angle. Instead, this multifactorial relationship depends on patterns of motoneuronal activity in relation to body configuration and body orientation to gravity.

Such complexity led to the widespread belief that spindle receptors were primarily associated with reflexes related to postural control rather than the perception of posture per se. For example, when a neurologist taps on a patient's tendon, such as the patellar tendon of the knee to elicit a knee jerk response, he or she is testing the integrity of the muscle spindle system. Tapping the tendon puts tension on the equatorial region exciting spindle primaries and secondaries. Their activity is conveyed to the spinal cord and has multiple consequences, including activation of the alpha motoneurons of the muscle that is tapped, causing a brief contraction coupled with inhibition of the motoneurons of the antagonist of the tapped muscle. In addition, the muscle spindle signals are conveyed upstream to the brain in the rapidly conducting dorsal columns of the spinal cord. Interestingly, if the patient strongly contracts his or her hands, the same tap to the knee will elicit a much stronger reflexive knee jerk response because the spindle sensitivity is being heightened by this maneuver.

Observations such as this made it seem unlikely that spindle signals could be related to position sense because the same or different activity levels could be associated with many different joint configurations. However, an important discovery in 1972 by Peter Matthews and his colleagues at Oxford University in England led to a revolutionary new perspective. It had been known for some time that if one vibrated the surface of a skeletal muscle with a physiotherapy vibrator, then the mechanical stimulation would cause the spindle receptors to fire. Their firing led to a contraction of the vibrated muscle—sort of like a continuous knee jerk response—that was known as a tonic vibration reflex or TVR. For example, vibration of the biceps muscle of the upper arm would result in reflexive flexion of the forearm. Matthews and his group discovered that if they resisted the flexion of the unseen forearm, the individual would experience illusory extension of the stationary forearm. Figure 1 in the Proprioception entry in this encyclopedia illustrates the illusion. The illusion of displacement and motion was completely compelling. Some people even experienced displacement of their forearm into anatomically impossible positions of hyperextension. Further studies led to the finding that abnormally high levels of spindle activity induced by vibration are always interpreted as the vibrated muscle being longer than it actually is, with this lengthening being attributed to a changed joint angle. Around the same time, joint replacement surgery became commonplace and a curious fact emerged. Patients with complete joint replacement retained an accurate kinesthetic sense despite the destruction of all the sensory endings associated with the joint and periosteum. This finding strengthened the case for muscle spindles being important for kinesthesia.

It turns out that for the spindle signals to be interpreted, the entire postural and motor context is critical. For example, if one strongly contracts a muscle that is being vibrated, the vibration will not elicit an illusory motion of the limb because the spindle discharge is already maximal. Thus, spindle signals have to be evaluated in the context of the current patterns of alpha and gamma motoneuronal activation. In addition, the ongoing configuration of the body relative to gravity is critical because this influences the relative loading of the muscles as it does when objects are being held. Golgi tendon organs are needed in the process of interpreting spindle discharges because they are accurate gauges of muscle tension. Together, such information is critical in coordination with a readout of descending motor commands. The ability to monitor such efferent commands is also key for the control of eye movements.

Skin has also been found to be a significant contributor to kinesthetic sensations, including the perception of an object's weight. Kinesthetic sensitivity is diminished when the skin of the toes or fingers is anesthetized. By contrast, anesthetizing the receptors within the joint capsule has little effect on sensitivity. Notice that when the elbow is flexed the skin on the back of the elbow is stretched and the skin in the crook of the elbow on the forearm and upper arm are in contact. This stretching and overlap conveys information about joint angle. The situation is similar with the fingers and toes, and their contact with adjacent digits also conveys interpretable information about their spatial configuration.

A popular demonstration of the role of the skin in kinesthesia concerns an anatomical peculiarity of the middle finger. If you strongly flex your middle finger while holding the other fingers of your hand strongly extended, you cannot voluntarily move the tip of your middle finger because its tendon is effectively disengaged. You can freely move the tip back and forth with your other hand without feeling any resistance. Yet, you can sense its ongoing position because of receptors in the skin although its position is much less accurately gauged than when the muscle and tendons are fully engaged. Such observations mean that skin receptors and muscle spindles (and Golgi tendon organs) work synergistically to provide kinesthetic information. Another important function of skin may be in potentiating muscle spindle activity. Experiments on the stretch reflex of the thumb that is mediated by muscle spindle receptors have shown that the strength of the reflex is diminished if the skin of the thumb is anesthetized. Such observations underscore the convergence of skin and muscle spindle signals in contributing to kinesthetic performance.

Phantom Limbs

Many amputees experience phantom limbs—it feels as if their missing limbs are still present. If an amputee moves the stump of the limb, it feels as if the phantom moves with the stump. Phantom limbs can be experienced with virtually all the same properties as normal limbs, such as sweating, heaviness, or temperature sensations. Importantly, many amputees can also "move'" their phantom limbs by an effort of will, for example, by flexing their phantom forearm or kicking with their phantom leg. Some phantom limbs can also be painful as when the phantom limb is a tightly clasped hand in which the "fingers" are painfully digging into the "palm." Temporary relief may be gained by voluntarily unclenching the phantom hand.

Attempts have been made to abolish phantom limbs by cutting the sensory nerves in the stump of the severed limb or blocking the activity of the nerves in the stump. These attempts usually fail completely. This failure means that the signals representing the phantom limb are more centrally generated. Importantly, however, it has been found that when an amputee can voluntarily move or control a phantom limb, there is activation of muscles in the stump that would normally have moved the missing limb. When there is no electrical activity in those muscles, movement of the phantom limb does not occur. This restriction makes it likely that sensory signals from the Golgi tendon organs and muscle spindles in the stump muscle interpreted in relation to the command signals to the muscle result in experienced movement of the phantom limb.

Sense of Effort and of Fatigue

An important component of kinesthesia is the sense of effort associated with moving a limb. One of the most prominent features of strokes that lead to impairments of movement is that the patient during the initial phase of paralysis is unable to move the impaired appendage. The patient wants to move it and attempts or wills to move it but feels no sense of effort or exertion. As the paralysis subsides and partial movement capability returns, it feels to the patient as if immense effort is involved in moving the limb and the appendage feels extremely heavy. Gradually, as normal movement returns, the sense of effort required to move the limb seems more normal and the limb no longer feels so heavy.

The sense of increased movement effort after cerebral strokes is associated as well with a variety of other disorders, including amyotrophic lateral sclerosis (Lou Gehrig's disease), polio, multiple sclerosis, lower motoneuron injury, and spinal cord injuries. Disuse—which affects motoneuronal and muscular response characteristics—also increases fatigue. A common feature of situations involving an increased sense of effort, heaviness, and fatigue is that the cortical motor output has to be increased above pre-injury or pre-disease levels to bring about movements. Sensory feedback does have a role, too. For example, when a normal person is holding a weight by supporting it through biceps activity, if the biceps is vibrated to elicit a TVR, it will seem as if less effort is necessary to support the load. By contrast, if the antagonist muscle, the triceps, is vibrated, the weight will seem heavier and it will seem as if more effort is required to support it. Even in cases where the tension is exactly the same, vibration of the agonist muscle or of the antagonist muscle gives the impression of lesser or greater effort and muscle force being exerted, respectively.

It is not fully understood why an increased sense of effort is experienced in the conditions and situations described. What is known is that projections to the cerebral cortex of muscle spindle and joint-related signals are primarily to the parietal lobe and that muscle spindle projections to the cerebellum are dense. These regions have reciprocal connections with the motor cortex where motoneuronal innervation of the muscles whose spindles are represented is originating. Thus, it is likely that many of the same computations involved in generating the sense of limb position and change in limb position are involved in creating the sense of effort.

James R. Lackner

See also Action and Vision; Body Perception; Corollary Discharge; Phantom Limb; Proprioception; Vestibular System

Further Readings

Brodal, A. (1973). Self-observations and neuro-anatomical considerations after a stroke. *Brain, 96,* 675–694.

Edin, B. B., & Johansson, N. (1995). Skin strain patterns provide kinaesthetic information to the human central nervous system. *Journal of Physiology, 487(Part 1),* 243–251.

Gandevia, S. C., Rowell, S. B., Rowell, L. E., & Shepard, J. T. (1996). Kinesthesia: Roles for afferent signals and motor commands. In L. B. Rowell (Ed.), *Handbook of physiology. Exercise: Regulation and integration of multiple systems. Neural control of movement, Sect. 12, Pt.1.* (pp. 128–172). Bethesda, MD: American Physiological Society.

Goodwin, G. M., McCloskey, D. I., & Matthews, P. B. (1972). The contribution of muscle afferents to kinaesthesia shown by vibration-induced illusions of movement and by the effects of paralysing joint afferents. *Brain, 95*(4), 705–748.

Jones, L. A., & Lederman, S. J. (2006). *Human hand function.* Oxford, UK: Oxford University Press.

Knoblich, G., Thornton, I. M., Grosjean, M., & Shiffar, M. (2006). *Human body perception from the inside out.* Oxford, UK: Oxford University Press.

Matthews, P. B. C. (1988). Proprioceptors and their contribution to somatosensory mapping: Complex messages require complex processing. *Canadian Journal of Physiology and Pharmacology, 66*(4), 430–438.

McCloskey, D. I., (1978). Kinesthetic sensibility. *Physiological Reviews, 58*(4), 763–820.

Sherrington, C. S. (1907). On the proprioceptive system, especially in its reflex aspect. *Brain, 29,* 467–485.

L

LANGUAGE

All human communities have, and use, *language*. Language allows humans to refer to objects, properties, actions, abstract entities, and other aspects of the world, and to convey and retrieve thoughts in a way that seems both fast and effortless. Both in its complexity and internal structure and in its expressive power, human language is well beyond any communicative system available to nonhumans. This entry surveys some basic empirical evidence and theorizing about the nature and properties of human language, the way language is produced and understood, and the way language is acquired by children.

The Nature of Language

Even though there are about 4,000 languages in the world today, they all share major design features that characterize the human faculty of language in general. One such design feature is creativity: Speakers of a language can produce and understand sentences that they have never uttered or heard before. For instance, it is possible to understand the meaning of the sentence *Napoleon never went to the moon* (and agree that the sentence expresses something true) without ever having encountered this sentence before. This shows that the human language ability does not rest on memorizing and storing linguistic strings but, rather, involves drawing on the finite number of items in one's vocabulary and combining them in novel and systematic ways to form a potentially infinite number of new sentences.

A second design feature of language is structure: Structural principles and rules constrain the kind of sentences that speakers can generate. Language involves several interconnected levels, from sound to syntax to meaning, and each level is governed by a specialized set of rules. For instance, one rule governing the sound structure of English specifies that the sequence *tl* cannot be used in the beginning of a new word (this is why *Tlong* is an unlikely name for a new cartoon character but *Klong* is not). Another kind of rule specifies the relative position of adjectives and nouns (*red umbrella* is an acceptable English phrase but *umbrella red* is not). A major goal for formal theories of language has been to discover the full set of principles underlying human languages. Following the work of Noam Chomsky, a further goal for many theorists has been to identify a core set of such basic principles (otherwise known as universal grammar) that can serve as the innate basis for all human languages.

The Production and Comprehension of Language

Humans can comfortably produce speech at the rate of four words per second, and comprehension follows the speed of production. Even though speaking and understanding speech seem effortless, both processes are supported by complex cognitive machinery.

Speaking involves multiple overlapping stages beginning with the intention to convey a message and ending with the formulation and execution of a sentence encoding that message. Speakers work in a top-down way, first deciding what they want to express, then choosing words and structures to communicate their message, and finally programming the implementation of the linguistic stimulus on the sound level so they can articulate it. The least well understood of these phases is the process of message preparation: Although this stage includes nonlinguistic apprehension of events and objects (and therefore interfaces with perceptual/conceptual representations of the world), the precise form of these nonlinguistic representations remains elusive.

More is known about the processes underlying the selection and assembly of words and sounds of a sentence. Evidence from several laboratories points to some degree of independence between the word- and sound-combining levels of speech planning: For instance, the tip-of-the-tongue phenomenon (the familiar sense that we cannot retrieve a word that we know) reveals that speakers can access information about the grammatical class of a word (e.g., whether it is a noun or verb) without accessing information about how the word sounds. Further evidence for the systems underlying language production comes from speech errors (or slips of the tongue). Such errors reveal rule-like constraints in how words or sounds are arranged when we prepare speech. For instance, when words switch places during a slip of the tongue, they almost always exchange with other words from the same grammatical class (nouns, verbs, etc.; e.g., *swimmers sink* becomes *swimmers drown*). When sounds switch places, they almost always exchange with other sounds of the same class of linguistic sounds (vowels or consonants; e.g., *snow flurries* becomes *flow snurries*). Thus, even speech errors seem to involve principled choices over abstract representations on the word or sound level—thereby offering indirect evidence for similar mechanisms in regular, error-free production. Such errors further demonstrate that speaking involves the active construction of utterances "on the fly" from smaller linguistic units.

Language comprehension (or *parsing,* as it is often called), unlike production, might appear to work in a bottom-up way: hearers are often thought to start with the sounds they hear, then identify the words and group them into structures, and proceed to infer what the speaker meant by a sentence. However, evidence indicates top-down influences on parsing because hearers may bring real-world knowledge or expectations about what speakers are trying to do to bear on the ongoing interpretation of an utterance. For instance, when people are placed at a table with several objects on it and are asked *Pick up the candle,* they move their eye gaze to the candle before they reach to perform the requested action. These spontaneous eye movements are a good measure of people's interpretation of incoming linguistic material. It turns out that hearers start looking at the candle approximately 50 milliseconds (ms) before the end of the word *candle.* But if there is candy on the table along with the candle, eye movements to the candle start only 30 ms after the end of the *candle.* This shows that auditory information as well as contextual information (the specific objects in a scene) affects word identification.

Similar evidence exists for structure identification. Consider a sentence that begins as follows: *Put the apple on the towel . . .* This may mean either that the apple should be placed on the towel, or that the apple that is on the towel should be placed in some location (to be specified in the rest of the sentence). In the absence of context, listeners prefer the first interpretation (e.g., if the sentence continues with . . . *into the box,* listeners become temporarily confused). But when placed at a table with two apples, one of which is on a towel and the other on a napkin, listeners' eye movements show that they access the second interpretation from the beginning (sentence continuations with . . . *into the box* cause no comprehension problems). During the interpretation of a sentence, then, syntactic, lexical, and contextual factors are rapidly integrated and affect the process of grouping words into linguistic structures. This example illustrates how the context provided by perception of a scene can play a role in this process of linguistic grouping.

As this and other experimental evidence shows, language comprehension is incremental: When hearers encounter a sentence, they begin to incorporate incoming words into a growing, richly structured representation of the sentence (rather than store them as a list and wait for the end of the sentence to recover the structure). Furthermore, hearers attempt to connect this representation of the sentence to the world around them. The precise way in which syntactic, lexical, and contextual factors affect parsing, as well as the way in which the language comprehension processes coordinate with the systems underlying language production, remain major questions in the psychology of language.

The Acquisition of Language

All human beings, under normal rearing circumstances, acquire language. Language learning begins at birth, if not earlier. Newborns are able to discriminate between possible sounds of human languages, even when these sounds do not belong to their native languages. Around the first year, as children begin to acquire the sound system of their native tongues, the ability to distinguish between foreign sounds is mostly lost. Children can understand certain words as early as 9 months, and they start producing words around their first birthday. First words (mostly, names for objects or individuals) are used in isolation and later in simple (two-word) sentences. Around the age of 2 or 3, grammatical markers (such as *-ed* and *-ing* in English) appear and the rate and diversity of vocabulary grows (with verbs, adjectives, and other terms being added). Between the ages of 3 and 5, children's sentences increase in length and complexity; a typical 5-year-old already knows about 10 to 15 thousand words.

The acquisition of language seems to take place fast, effortlessly, accurately, and mostly without explicit instruction from adults. Children can acquire language even in societies where no language is directed to infants before infants themselves speak. How is this feat accomplished? Part of the answer is that language learning requires the discovery of regularities on multiple levels of the linguistic input, and evidence shows that even very young infants are sensitive to linguistic patterns (such as recurring sound sequences). However, most theorists agree that children's search for patterns in the linguistic input has to be guided by strong biological influences. Several arguments support this position. First, universal properties of language surface in all the languages that have been studied so far. For example, linguistic sentences always include a subject (even though in some languages such as Greek, the subject is not overtly pronounced but is understood from the form of verb: *Efiga htes* = "[I]-left yesterday"). Second, as described, language acquisition seems to proceed universally through the same stages for all learners despite differences in culture, socioeconomic level, parenting style, motivation, and other factors in the learners' environment. Additionally, linguistic abilities seem to be specialized and often dissociate from general-cognitive abilities in pathology. For instance, children with Williams syndrome, whose IQs are below normal, have intact language-learning capacities; children with specific language impairment are characterized by deficits in the time-line and nature of their language learning, even though their IQs are within the normal range.

Some of the most compelling evidence for humans' biological preparedness for language comes from the resilience of language learning given absent or degraded input. Children who are not exposed to conventional language (e.g., deaf children growing up among hearing adults without access to sign language) have been known to invent spontaneous gesture systems to communicate. Crucially, the properties of these systems seem to be similar to those of early speech (children begin with simple gestures, then produce two-gesture combinations and later more complex "sentences"). Similarly, children exposed to improvised and irregular language-like systems (*pidgins*) constructed by members of different language communities who find themselves living in the same environment transform these imperfect systems, as they learn them, into rule-governed languages (*creoles*). Finally, children whose parents are non-native speakers of either a spoken or a signed language typically regularize the incorrect linguistic forms they are exposed to. In all these cases, children go beyond the information they encounter in the input—in a sense, they create, rather than simply acquire, language.

Further support for the conclusion that learners contribute substantially to language growth comes from studying how changes in a learner's mental preparedness affect language acquisition. Evidence from abandoned and neglected children suggests that the age at which these children were found had a profound effect on whether their language ability could be restored. Experimental data from sign language show that late exposure to a first (signed) language has deep negative effects on learning (similar effects of late exposure hold for second language acquisition). Such critical period effects demonstrate that the maturational state of the learner's brain is crucial for the attainment of a language system.

Several questions remain open about the mechanisms underlying language learning. One issue is whether specific aspects of language acquisition should be attributed to language-specific versus general-purpose learning mechanisms. Another issue is whether children's native language can affect the way they think, and whether language is necessary or helpful for the development of human concepts.

Anna Papafragou

See also Aphasias; Audition: Cognitive Influences; Context Effects in Perception; Speech Perception; Top-Down and Bottom-Up Processing; Word Recognition

Further Readings

Baker, M. (2001). *The atoms of language.* Oxford, UK: Oxford University Press.

Bloom, P. (Ed.). (1993). *Language acquisition: Core readings.* Cambridge: MIT Press.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use.* New York: Praeger.

Clifton, C., Frazier, L., & Rayner, K. (Eds.). (1994). *Perspectives in sentence processing.* Hillsdale, NJ: Lawrence Erlbaum.

Gleitman, L., & Liberman, M. (Eds.). (1995). *An invitation to cognitive science: Vol. 1. Language.* Cambridge: MIT Press.

Pinker, S. (1995). *The language instinct.* London: Penguin Books.

Trueswell, J., & Tanenhaus, M. (Eds.). (2006). *Approaches to studying world-situated language use.* Cambridge: MIT Press.

LATERAL INHIBITION

Lateral inhibition, is a decrease in response in neurons that occurs when neighboring neurons become activated. For example, in Figure 1(a), a network of 10 excitatory neurons receiving information from visual space (such as neurons in the retina or later levels of the visual system) is intermingled with 9 inhibitory neurons. Activity in any one of the excitatory neurons can inhibit its neighbors indirectly by activating the inhibitory neurons that then inhibit their neighbors. When a stimulus (such as a bar of light or any other stimulus) excites a number of neurons in the network (in this case neurons 4e, 5e, 6e, and 7e), the effect of inhibition is to suppress the neurons just outside the edge of the bar (3e and 8e) because those neurons are inhibited but not excited. Further, because the neurons just inside the edges of the bar (4e and 7e) are excited by light and only inhibited by one neighbor, they are especially active. This leads to perceptual contrast enhancement at borders. Further research showed that lateral inhibition also applied to overlapping stimuli, and that its strength fell off with distance between the interacting stimuli.

Haldan Keffer Hartline won the Nobel Prize in 1967 for discovering lateral inhibition and its neural correlates. The first inhibitory circuit in the nervous system was found in the horseshoe crab (*Limulus polyphemus*). Here, an activated photoreceptor was inhibited when a laterally adjacent (or nearby) photoreceptor was also activated. Lateral inhibitory circuits are currently known to be ubiquitous to all sensory areas of the brain, and they play an important role in many sensory, cognitive, motor, affective, and limbic processes. The most common mechanism by which neurons suppress their neighbors is through the inhibitory neurotransmitter gamma-aminobutyric acid (GABA).

Hartline and his collaborator, Floyd Ratliff, went on to characterize the three components of a laterally inhibitory circuit: (1) Excitatory input and output—information input arrives at a given sensory area of the brain in the form of excitatory neural responses. Information output is sent to the next area(s) in the hierarchy also in the form of excitatory neural responses; (2) Self-inhibition—neurons that laterally inhibit their neighbors also inhibit themselves; (3) Lateral inhibition occurs as a function of excitatory activation—thus inhibition follows excitation in time.

The Role of Lateral Inhibition Through Time

In addition to its effects across space, lateral inhibition also leads to temporal effects over time. Let us now examine two neurons embedded within a lateral inhibitory network as a function of time (Figure 1b): one excitatory neuron (at times 1e

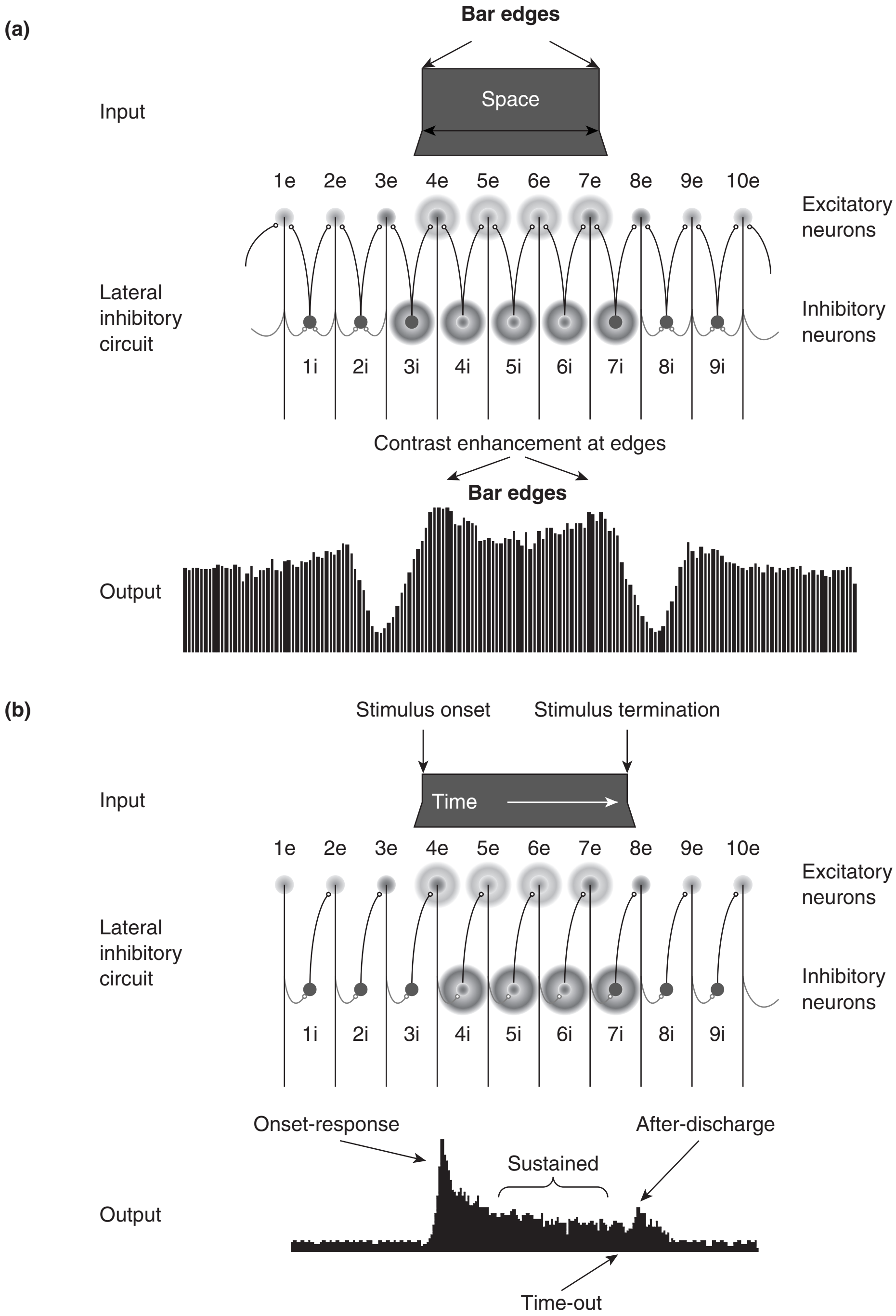


Figure 1 Spatial Lateral Inhibition Model

Notes: (a) A mammalian representation of the spatial lateral inhibition model originally proposed by Hartline and Ratliff. The excitatory neurons in the center of the upper row receive excitatory input from a visual stimulus. This excitation is transmitted laterally to the inhibitory neurons just outside the stimulus and within the area impinged on by the stimulus. The inhibitory interactions between excited neurons at the edges of stimuli and their non-excited neighbors results in apparent contrast enhancement at the borders of the stimulus. Output of each of the excitatory neurons is represented in action potentials per unit time at the bottom. (b) One excitatory and one inhibitory neuron taken from the spatial model in (a), now followed through a period of time in which the stimulus is off (times 1, 2, and 3), on (times 4, 5, 6, and 7), and then off (times 8, 9 and 10).

through 10e) and its connected inhibitory neuron (at times 1i through 9i). At times 1e, 2e, and 3e (before the stimulus is presented), there is no excitatory input, so the output remains flat. At time 4e (just after the stimulus, such as a bar of light, is presented), the neuron is excited, causing an onset-response. This leads to the activation of the inhibitory neuron at time 4i, after a slight delay. The inhibitory neuron then feeds back on the excitatory neuron and causes its activity to be suppressed at time 5e. This state of excitatory-inhibitory equilibrium is called the sustained period, which continues through time 7e, after which the stimulus is extinguished. Despite the stimulus having been terminated, the neuron at time 7i is nevertheless activated by the excitatory neuron at time 7e because of the delayed effect of inhibition. Thus, the excitatory neuron at time 8e is inhibited while not being excited by visual input, and so is in a state of deep suppression called the time-out period, which in turn causes the inhibitory neuron at time 8i to be deeply suppressed because of lack of input. The excitatory neuron at time 9e then exhibits a disinhibitory rebound called the after-discharge because of the lack of baseline inhibition (even though there is no excitatory input).

Therefore, just as lateral inhibition causes neurons to respond strongly to spatial borders of stimuli, it also makes them respond strongly to temporal borders (the onsets and terminations) of stimuli. The perceptual result of this is contrast enhancement at the temporal borders of stimuli.

Stephen L. Macknik
and Susana Martinez-Conde

See also Contrast Enhancement at Borders; Vision: Temporal Factors; Visual Illusions

Further Readings

Hartline, H. K. (1939). Excitation and inhibition of the "off" response in the vertebrate optic nerve fibers. *American Journal of Physiology, 126,* 527.

Macknik, S. L., & Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nature Neuroscience, 1*(2), 144–149.

Macknik, S. L., & Martinez-Conde, S. (2004). The spatial and temporal effects of lateral inhibitory networks and their relevance to the visibility of spatiotemporal edges. *Neurocomputing, 58–60C,* 775–782.

Ratliff, F. (1965). *MACH BANDS: Quantitative studies on neural networks in the retina.* San Francisco: Holden-Day.

LAZY EYE

See Amblyopia

LIGHT MEASUREMENT

Light provides humans with stereoscopic images of the world at all distances. This entry considers measuring light as it affects human vision. Humans are most sensitive to electromagnetic radiation in a small window between 400 and 700 nanometer (nm) wavelengths (λ). This response begins with four types of retinal receptors. After sitting in a light-free room for an hour, humans report seeing light with only four to six photons. Snow on a mountaintop sends to the eye 100 million times more photons. Although photographic film has a fixed, unique response to the number of photons/area, human visual appearance has a complex spatial relationship to the light on the retina. Nevertheless, light measurement in psychophysics is important: first, to accurately describe the display presented to observers, so others can reproduce the experiment; and second, to describe the light array, pixel (picture element) by pixel, of the entire field of view, as input for computer appearance models.

Radiometry

Radiometry, a part of physics, describes standards for measuring electromagnetic radiation. Photons with different wavelengths have different energies. From Planck's law, we can calculate that a single 555-nm photon's energy equals 3.6×10^{-19} joules (or, watt × seconds).

Irradiance is the measure of the energy from the number of photons continuously falling on an area. Irradiance meters measure (watts/centimeter $[\text{cm}]^2$)

using photosensitive electronic elements behind a diffuser integrating the light falling on the meter. Photographers use similar incident light meters to measure the light falling on the scene by standing in front of the subject and aiming the meter at the light source.

Radiance is the measurement of light continuously coming from a particular object to a particular point in space. Ansel Adams' zone system described techniques to capture and print the entire range of all scenes. He measured the light from clouds, sky, rocks, trees, and shadows using a spot photometer, a camera-like instrument that forms an image of a small scene segment on the light-sensitive element. Such telephotometers measure photons per area per angle (watts/[cm^2 × steradian]). By replacing the average of the entire scene with individual values for maximum light, minimum light, and objects of interest, the photographer can control the image-making processes to render his or her visualization of the scene.

Photometry

Photometry, a part of psychophysics, is radiometry adjusted for the wavelength sensitivity standards for human vision. The photopic luminosity standard (V_λ) is the nominal sensitivity of the eye to different wavelengths. The peak sensitivity is at 555 nm (100%); half-height (50%) at 510 and 610 nm; (3%) at 446 and 672 nm. The candela/meter^2 (cd/m^2), the standard of luminous radiance, is 1/683 watts per area per angle at 555 nm. This standard is the basis of lumens per watt used in evaluating the electrical efficiency of lamps (Lumen = [candela × angle]).

To convert the radiometric units (physics) *irradiance* and *radiance* to photometric units (psychophysics) of *illuminance* and *luminance,* we multiply the radiometric values at each wavelength by the *Commission Internationale de l'Eclairage* (CIE) standard V_λ curve value of those wavelengths. Illuminance is the integral of (irradiance$_\lambda$ × V_λ); luminance is the integral (radiance$_\lambda$ × V_λ). Before 1930, the term *brightness* was defined as both the measurement of light and the measurement of appearance. The term *luminance* was introduced to measure light and to reduce the confusion caused by the two usages of brightness with different meanings.

Color Measurement

The V_λ achromatic standard converts radiometric light units to photometric measures of human responsiveness. Color requires colorimetric standards. V_λ was defined by measuring human response to certain stimuli. Colorimetric standards do the same, by asking observers different questions. Two important experiments are used in measuring color.

First, color matching functions (CMF) ask observers to match a single wavelength in one half of a circular field by adjusting the amounts of red, green, and blue (RGB) light in the other half of the field The 1931 CIE standard describes three spectral responses that are consistent with a large pool of color matching data documented by Andrew Stockman and Lindsay Sharpe. Just as with V_λ, we calculate the integral of ($X = \text{radiance}_\lambda \times \bar{x}_\lambda$), ($Y = \text{radiance}_\lambda \times \bar{y}_\lambda$), and ($Z = \text{radiance}_\lambda \times \bar{z}_\lambda$). These XYZ values are effective in calculating when two stimuli will match. However, as Gunter Wyszecki warned us, these XYZ values do not provide any direct clues for the color appearance of these matching stimuli. Additional variables, such as the other colors in the field of view, the state of adaptation of the receptors, and spatial relationships in each scene, determine the color appearance. To measure appearance, one needs to match a test color to a library of standard colors in a constant complex scene.

The second data set in standardizing color is the *Munsell Book of Color* and other uniform color spaces (UCS). Here, observers equated the visual distance between colored samples in hue, chroma, and lightness. Think of all possible colors as a three-dimensional (3-D) crystal structure. These experiments assigned colored papers to each uniformly spaced lattice point in the crystal. Unlike CMFs that are measured in a no-light surround, Munsell designations are chosen for viewing in complex scenes. Furthermore, they measure the complete range of appearances in a 3-D color space. Physical measurements of the papers selected in the Munsell book provide a measurement of light consistent with that of color appearance in real scenes.

When we plot the equally spaced in appearance Munsell Color chips in the colorimetric XYZ space, we find unequal XYZ distances. To convert XYZ values to a UCS, we need a complex, nonlinear transform. In other words, color matching standards do not have a simple correlation with appearance.

Digital Imaging

With digital imaging, one can capture and display scenes effortlessly. With added efficiency and automatic image processing, it is essential to remember the critical importance of calibration measurements of displays used in psychology experiments. Digital values for each pixel in a photograph of a scene are not equal to, nor proportional to, the array of radiances from the scene. The automatic image processing in all cameras and all displays apply complex transforms of scene radiances to achieve pleasing images. As well, digits in computer memory are not proportional to displayed radiances. One must carefully calibrate all electronic devices, over the entire 3-D color space, to understand the relationship of RGB digits and radiance. The outermost regions of the 3-D color space—including whites, blacks, and all high-chroma colors—are often of importance and interest in psychology experiments. Printers and displays require careful calibration of these near-gamut-boundary colors because devices have the least predictable control of light stimuli for these colors.

In photography, there is a simple relationship between the amount of light (number of photons at each wavelength) and the response of the camera sensor. The same number of photons will initiate the same response everywhere in the image, for every image, every time. In psychology, we observe the complex responses of the human visual system to the number of photons. We observe that the same number of photons do not initiate the same response in the same image (simultaneous contrast), for every image (scene dependence), and every time (adaptation). Photometric standards are helpful in describing stimuli and improving communication between authors. In psychology, however, there is no simple, precise, determinative relationship between measured light and appearance.

John J. McCann

See also Color Perception; Vision; Visual Stimuli

Further Readings

Committee on Colorimetry, Optical Society of America. (1953). *The science of color.* New York: Thomas Y. Crowell.

McCann, J. J. (2005). Rendering high-dynamic range images: Algorithms that mimic human vision. *Proceedings of Advanced Maui Optical and Space Surveillance (AMOS) Technical Conference, Maui* (pp. 19–28).

McCann, J. J. (2007). Art, science, and appearance in HDR images. *Journal for the Society of Information Display, 15*(9), 709–719.

McCann, J. J., & Rizzi, A. (2007). Camera and visual veiling glare in HDR images. *Journal for the Society of Information Display, 15*(9), 721–730.

Stockman, A., & Sharpe, L. T. (n.d.). Color & vision database. Retrieved from http://cvision.ucsd.edu/

Wyszecki, G. (1973). Colorimetry. In R. Eynard (Ed.), *Color theory and imaging systems, Society of Photographic Scientists and Engineers* (pp. 24–49). Washington, DC: Society of Photographic Scientists and Engineers.

Wyszecki, G., & Stiles, W. S. (1967). *Color science: Concepts and methods, quantitative data and formulae.* New York: Wiley.

LIGHTNESS CONSTANCY

Lightness constancy is the ability to perceive black, gray, and white surfaces correctly regardless of illumination level. Lightness is the perceptual equivalent to the physical shade of gray of a surface, technically known as reflectance. A white surface reflects about 90% of the light it receives whereas black reflects only about 3%. Thus, a white paper will reflect 30 times as much light as a black paper lying next to it in the same illumination. This reflected light enters the eye when we look at the papers, and the light-sensitive rods and cones are stimulated proportionately. This seems to explain our perception of black and white, except that the intensity of illumination is seldom uniform in our world. Illumination intensity can vary by a factor of more than a million-to-one, so a black surface in sunlight can easily reflect more light than a white surface in shadow does. A black Labrador retriever lying in the sunlight might reflect more light than does a white rabbit sitting in the shade of a tree. Thus, by itself, the light that comes to the eye from a black, gray, or white surface, unlike that from a colored surface, contains no hint of the shade of gray from which it comes. Indeed any intensity of light can come from any shade of gray.

The only way a visual system (human or robotic) can determine the shade of gray of a surface in its field of view is by using the surrounding context. But researchers do not agree about exactly how that context is used. No robot can recognize whether a surface is black or white. Any computer can encode and store a digital image, but determining the shade of gray of an object within that image is currently beyond our reach because no one is able yet to write the software.

Although the absolute intensity of light (called luminance) reaching the eye from an object reveals nothing about the reflectance of that object, relative luminance is more useful. A square patch of light of fixed luminance in the middle of a computer monitor can be made to appear as any shade between black and white merely by changing the luminance of a homogeneous background that surrounds it. As long as the square is brighter than its surround, it will look white. As soon as the surround is made slightly brighter than the square, the square begins to appear slightly gray, and as the luminance of the surround continues to increase, the square will appear to become darker gray, finally appearing completely black when the luminance of the surround is about 30 times greater than the luminance of the square.

Our sensory apparatus is specialized to detect relative luminance. The human eye is in constant motion, including 30 to 150 random jiggling motions per second. If these motions are prevented, say by paralyzing the eye muscles, the visual scene disappears. Apparently, it is critical that the rod and cone photoreceptors in the eye be able to sweep across the change in luminance that occurs at each boundary in the image. If the visual field is made completely homogeneous, say by placing an observer's head inside a dome so that its homogeneous interior surface fills the entire visual field, no shade of gray will be seen. Indeed no surface will be seen—the observer perceives only an infinite fog.

Subtle changes in context can have a profound effect on perceived gray shade, especially if they influence the perceived spatial position of a surface. Figure 1 illustrates a simple scene in which a wall containing a doorway separates a dimly lit near room from a brightly lit far room seen through the doorway. The square labeled 30 will appear almost white when it appears to lie in the near room, as shown on the left. But it will appear almost black when it appears to lie in the far room, as shown on the right, even though the relative luminance values are the same in the two cases.

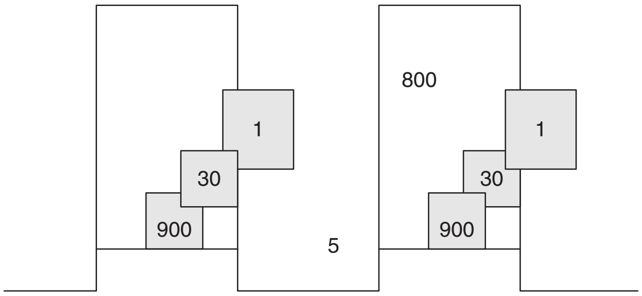


Figure 1 Perceived Lightness Depends on Perceived Depth

Notes: This line drawing illustrates a doorway in a near, dimly lit room, that reveals part of a far, brightly lit room. The numbers indicate the relative amounts of light (luminance) reflected from each surface. On the left, the square labeled 30 appears to be located in the near room and it appears almost white. On the right, it appears to be located in the far room and it appears almost black (Note: All squares are shown as gray in the figure, but change their appearance as indicated, depending on where they are perceived to be located).

Ultimately, relative luminance can specify only relative lightness, but not a specific gray shade. For example, a white surface next to a middle gray surface can produce a relative luminance of 5:1 in the retinal image. But so can a middle gray surface next to a black surface. A fundamental problem for theories of lightness, called the anchoring problem, concerns how relative luminance values at boundaries in the image are processed to yield the specific shades of gray that we see. We can give the visual system only a single value of relative luminance, say 5:1, by placing an observer's head inside a dome that is divided into two homogeneous halves by a sharp boundary. In this case, the observer will see the 5 as white and the 1 as middle gray, even if the actual surfaces are painted middle gray and black. Apparently, the fundamental anchoring rule is this: The highest luminance in a scene appears white.

For simple images such as these, we can write the rules used by the visual system to determine the gray shades we actually see. But for more complex images, produced by the scenes of everyday life, there is no consensus, only competing theories. These fall roughly into three classes: low-level, high-level, and mid-level. According to high-level, or cognitive, theories, lightness constancy results from a thought-like inferential process that includes learning. These theories are often intuitively appealing, but are criticized as being too vague. Fish, chickens, and human infants show good lightness constancy, which further undermines the cognitive approach, suggesting that lightness constancy is hardwired from birth.

Low-level theories are closely tied to our current understanding of retinal physiology. These tend to be simpler theories and highly mathematical, but they are criticized for failing to capture important aspects of lightness perception, such as the influence of depth perception illustrated in Figure 1. Mid-level theories attempt to combine the advantages of both high- and low-level theories. They are more sophisticated than low-level theories but more concrete than high-level theories.

Lightness constancy represents a remarkable achievement by the human visual system. With time, and additional research, we will understand how it works.

Alan Gilchrist

See also Color Constancy; Constancy; Contrast Perception; Depth Perception in Pictures/Film

Further Readings

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science, 262,* 2042–2044.

Gilchrist, A. (1979). The perception of surface blacks and whites. *Scientific American, 240,* 112–123.

Gilchrist, A. (2006). *Seeing black and white.* New York: Oxford University Press.

Wallach, H. (1963). The perception of neutral colors. *Scientific American, 208,* 107–116.

LIGHTNING AND THUNDER

Among the most spectacular atmospheric phenomena are lightning and thunder, which are described in relation to perception in this entry. Lightning and thunder involve making light, sounds, and smells from the discharge of static electricity by clouds of water or ice, of dust or sand, and of ash. Discharges include *lightning, St. Elmo's fire, sprites, blue jets,* and *elves.* The latter three are lightning from the tops of clouds into the upper atmosphere and are difficult to see from the ground. But lightning, a blue-white arc of light about 10 times brighter than the sun, makes thunder with a peak loudness of 170 decibels (dB) and produces a faint odor of ozone. Lightning can blind, deafen, disable, and kill. St. Elmo's fire, a dim, blue-white steady glow from sharp objects, is visible at night, audible as a hiss, and produces the same odor.

These phenomena involve electrical charge from clouds in two phases. First, an electrical field becomes sufficient to produce *coronal discharge* from movement of charged particles, ions and electrons, through air. St. Elmo's fire consists only of this first phase. If the area of ionization grows, it becomes the first phase of lightning.

In lightning, massive electrical potential builds up in a cloud, allowing electrons and ions to move along narrow channels between a cloud and other clouds (inter-cloud; 75%) or the ground (cloud-ground; 25%). The charged particles move in a zigzagged, branched path, taking about 35 milliseconds (ms); this phase is usually invisible. But once a path is completed, the second phase occurs: a 20,000-ampere (amp) current, the *return stroke,* flows for about 0.1 ms at 0.4 times the speed of light to the cloud along a *lightning channel* about 1.3 centimeters (cm) wide and 1.6 to 7 kilometers (km) long. This produces all wavelengths of the electromagnetic spectrum, including light. Within it, the current raises the temperature of the air, essentially instantaneously, to about 30,000°C producing acoustic shock waves. These decay into sound waves after traveling a few meters and are heard as thunder.

After a first return stroke, another can occur along the same charged path after about 50 ms. As many as 40 return strokes can occur in a single lightning; the average is 4, yielding *stroboscopic illumination.* Strong wind can move the charged path so that subsequent return strokes occur in appreciably different parts of the sky, yielding *apparent movement.*

When lightning is far from us, we see the light long before we hear its thunder. This is because the light reaches our eyes at the speed of light, essentially instantaneously, whereas the thunder reaches our ears at the speed of sound, about 0.33 km per second.

We see *sheet lightning* because the return stroke occurs behind other clouds, diffusing the light. Although the movement of most cloud-ground return strokes is upward, it is far too fast for us to see any movement. Nevertheless, people report lightning moving either downward or upward. This could occur if one part of a stroke were brighter than another, such as if one part is partially obscured by rain. The brighter part would appear to be earlier than the dimmer part, leading to apparent movement in the bright-dim direction. Occasionally, inter-cloud strokes can occur sequentially in different parts of the cloud and are visible as apparent movement.

Lightning looks much wider than the lightning channel. This is partially physical (visible in photographs) and partly perceptual. Physically, intense light is scattered in the retina or in film, affecting neighboring retinal cones or photographic molecules, enlarging the image. This is the same reason the sun looks like a blindingly bright fuzzy disk much wider than its 0.5-degree diameter. Perceptual reasons include a viewer's eye not likely being accurately focused on the lightning. Blurring of lightning's retinal image increases its width. We usually see lightning at night, reducing accuracy of focusing. At night, most people's eyes adopt a *resting focus* position, about 2 meters from the eye. Also the pupil is large, reducing the *depth of focus* of the eye. Because lightning is so far from the eye, focusing at 2 meters with large pupils causes serious blur of the image of the lightning. Finally, even if a viewer's eyes were accurately focused, a bright object looks larger than an identically sized dark object because of *irradiation,* an illusion named by Hermann von Helmholtz about 1867.

The quality and duration of the sounds of thunder depend on the distance from parts of the return stroke to the hearer. Thunder comprises acoustic energy from 0 hertz (Hz) to 2.5 kilohertz (kHz), with peak power around 20 to 60 Hz and falling off rapidly with frequency. Within 100 meters (m) of cloud-ground lightning, people hear a click from the positive streamer, then a short crack from the near parts of the return stroke, then rumbling from more distant parts of the return stroke. Within 200 m, the click is replaced by a sound like tearing cloth, probably from movement of charged particles from the ground toward the lightning channel (*positive streamers*). Near thunder sounds sharp, from high frequencies, and much shorter than its duration (described as "cracks" or "peals"), presumably from masking of later-arriving sounds. Distant thunder is a low rumble whose loudness waxes and wanes ("rolls" and "rumbles"), and is audible for as long as 41 seconds. From about 25 km, no sound can be heard although the lightning can still be seen, misleadingly called *heat lightning.*

The sound of distant thunder is affected by attenuation, refraction, reflection, and interference by the atmosphere and by echoes from the terrain. High frequencies are attenuated, leading to the low pitch of distant thunder. Air is denser and warmer near the ground than near the cloud. This refracts sounds to curve up from the ground toward the clouds. Distant thunder mainly comprises sounds from the high parts of lightning. But temperature inversion layers can reflect sounds back toward the ground; these late-arriving echoes increase the duration. Downwinds reduce refraction, leading to louder sounds; upwinds increase refraction, leading to fainter sounds. (This is why it is easier to hear ordinary sounds downwind than upwind.) Turbulence cells in the atmosphere can also alter the paths sounds take to an observer, reinforcing some sounds and canceling others.

Robert P. O'Shea

See also Atmospheric Phenomena; Audition: Loudness; Audition: Temporal Factors; Eye: Structure and Optics; Motion Perception

Further Readings

Rakov, V. A., & Uman, M. A. (2003). *Lightning: Physics and effects.* New York: Cambridge University Press.

Zavisa, J. (2000). How lightning works. Retrieved from http://www.howstuffworks.com/lightning.htm

LINEAR AND NONLINEAR SYSTEM ANALYSIS

Mathematics is critical for the analysis of perceptual and other brain processes because it represents the most powerful tool humans have developed for understanding and predicting the behavior of complex systems and their changing responses over time. The study of change over time is known as *dynamics*. *Linear systems* represent the simplest form of mathematical description of perceptual phenomena, and their characteristics are quite well understood. *Nonlinear systems* are much more complex and only partially understood. Fortunately, however, the types of nonlinearity characterizing the brain have been intensively studied for almost 50 years, so a substantial degree of understanding has been achieved. Most exciting is the observation that nonlinear systems can generate strikingly complex forms of response capable of explaining brain function, so this entry mainly focuses on such nonlinear behaviors.

Linear Systems

In studying perceptual systems, scientists typically present a stimulus (*S*) to the system and then measure the response (*R*) of the system. *S* might be as simple as a pattern of light and dark parallel bars or as complex as a human face. *R* might represent an observer's estimate of the contrast between the light and dark bars or a judgment of the direction in which the face is looking. If the system mediating these responses is linear, then it follows that

$$R = k \times S, \tag{1}$$

where *k* is a constant that defines the characteristics of the system. This is a static description, but no physiological system ever responds instantaneously to its stimulus. Therefore, it is necessary to introduce dynamics to represent the temporal development of the response as it approaches the final state described by Equation (1). If the change in response in a small period D*t* is represented by D*R*, then the rate of change of response will be D*R*/D*t* (technically a derivative in calculus). Thus, time can be introduced into Equation (1) by writing

$$\Delta R/_{\Delta t} = -R(t) + k \times S. \tag{2}$$

This is the most basic equation in all of linear systems theory. It states that the rate of change of the system response is equal to the difference between the current response *R*(*t*) and the final response state determined by the stimulus, *k* × *S*. When the system reaches equilibrium, where D*R*/D*t* = 0, the final response is again described by Equation (1). This is a linear equation because all of the individual terms (D*R*/D*t*, *R, S*) are linear; the actual response over time represents an exponential approach to equilibrium.

Multiple Components and Feedback

This description represents the simplest possible linear system in which the response *R* is determined directly and exclusively by the stimulus *S*. However, all perceptual systems involve multiple neurons that interact in a variety of ways. For example, *S* might trigger an initial response in neural receptors, and this response in turn might provide the input to a second sensory neuron that transmits the information to the brain. This would be called a feedforward network because the receptors feed their response to other neurons, which send their responses to the brain. A much more interesting case is that in which the receptors activate inhibitory neurons, which in turn send an inhibitory or negative response back to the receptors themselves. This is the case with the photoreceptors and inhibitory horizontal cells in the vertebrate retina. To describe this system, there must be an equation for each neuron. If the response of the receptor is *R*, and the response of the inhibitory cell is *I*, then Equation (2) can be generalized to

$$\begin{aligned} \Delta R/_{\Delta t} &= -R(t) + k \times S - I(t) \\ \Delta R/_{\Delta t} &= -I(t) + g \times R(t), \end{aligned} \tag{3}$$

where the constant *g* determines the strength of the inhibition. This D*I*/D*t* equation is just like Equation (2), except that the stimulation comes from the photoreceptor response *R*. The term –*I*(*t*) in the D*R*/D*t* equation describes the inhibitory or negative feedback onto the receptors.

Linear perceptual systems can have any number of components, and systems with 10 or more components are not uncommon. Furthermore, feedback can be positive between some components and negative between others. Indeed, the balance between positive and negative feedback determines how the system dynamics will evolve over time. With physiologically reasonable values of feedback constants such as *g*, a linear system will generally decay to an equilibrium such as described by

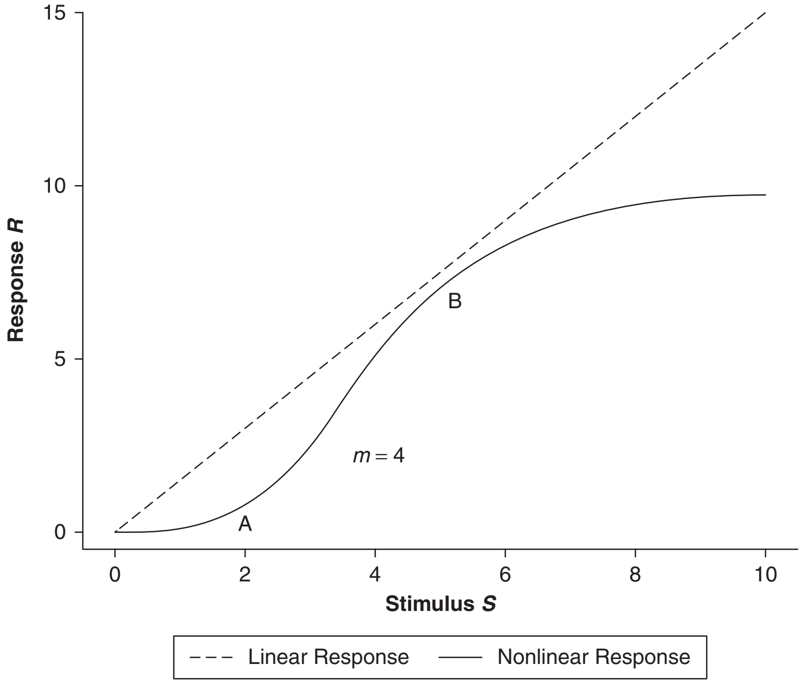


Figure 1 Examples of Linear (Dashed Line) and Nonlinear (Solid Line) Responses to a Stimulus *S*

Notes: The nonlinear response has a threshold at A, is almost linear between A and B, and saturates above B. The nonlinear response reaches half of its maximum value when $S = m = 4$.

Equation (1). The exact mathematical solution to Equation (3) and its generalizations can be written as a combination of exponential functions, sines, and cosines. However, much more complex behavior occurs when the system is nonlinear.

Nonlinear Systems

All perceptual and physiological systems are inherently nonlinear. A nonlinear system is any system incorporating at least one nonlinear function, such as R^2 or $\log(R)$. In consequence, there are vastly more (strictly, infinitely more) nonlinear systems than there are linear systems. As a simple but important example, responses R of many sensory neurons can be described by the nonlinear function

$$R = \frac{k \times S^2}{S^2 + m^2}. \tag{4}$$

The constant m determines the stimulus value S at which this nonlinear response function reaches half of k, its maximum value. Figure 1 compares this function, known as a Naka-Rushton function after the perceptual scientists who introduced it, with the linear function in Equation (1). This nonlinear response function incorporates an activation threshold because there is almost no response below the point marked A. The flattening of the curve above point B represents the fact that responses saturate at high stimulus levels because neurons cannot fire more than about 500 to 1,000 impulses per second. Between points A and B, the curve increases approximately linearly, so linear systems analysis can be used to gain insight into nonlinear systems when restricted to small regions of the nonlinear curve. The nonlinear curve in Figure 1 is often referred to as a sigmoid or S-shaped curve, and nonlinearities in the brain almost always involve such sigmoid nonlinearities.

With this background, Equation (3) can now be converted into a nonlinear form that provides much more accurate descriptions of neural systems:

$$\begin{aligned} {}^{\Delta R}\!/_{\Delta t} &= -R(t) + \frac{k \times S^2}{S^2 + m^2} - I \\ {}^{\Delta I}\!/_{\Delta t} &= -I(t) + \frac{g \times R^2}{R^2 + m^2}. \end{aligned} \tag{5}$$

This equation results from replacement of the two linear terms in Equation (3) with sigmoid non-linearities.

Equations of this type were introduced into brain modeling almost 40 years ago and are now known as the Wilson-Cowan equations for the scientists who first created them.

Inherently Nonlinear Dynamic Responses

Mutual excitation between two neurons provides a simple example of nonlinear behavior. Suppose that one neuron is excited by a vertical bar and the second is excited by red patches. If a red, vertical bar is viewed, both neurons will be excited, and they will in turn excite one another still further. If the mutual excitation is strong enough, the pair of neurons will continue to fire even after the stimulus disappears, thus forming a perceptual memory of the stimulus. Such perceptual memories have been shown to occur in the cortex of monkeys. Furthermore, weak stimulation of such positive feedback networks can provide a basis for understanding perceptual priming. The sigmoid nonlinearity is critical to this behavior for two reasons. First, the saturation above B in Figure 1 guarantees that the mutual excitation can only drive neural activity upward to a finite level. Second, the threshold at A in the figure makes it possible to turn the memory off by providing a negative or inhibitory stimulus to one of the neurons. In this manner, short-term memory in a nonlinear system can be erased to permit short-term storage of new memories. In a linear system, which has no thresholds, this is impossible.

Perceptual memory networks also exhibit another nonlinear dynamical phenomenon: hysteresis. A system with hysteresis can be in either of two states for a range of external stimulus levels, and the state it is currently in is determined by the history of stimulation. Probably the most familiar example of hysteresis is magnetic hysteresis. An iron nail, which is initially not magnetized, can be turned into an electromagnet by wrapping it with a spiral of wire and then passing electricity from a battery through the wire. After a minute or so of this, the battery can be disconnected and the nail removed from the spiral of wire, but the nail still retains its magnetism. The nail actually has a physical memory of its recent history of electrical stimulation, which results from positive feedback produced by alignment of the tiny magnetic fields of the individual iron atoms. In the perceptual memory example, the neural responses in the absence of any stimulus will remain at an elevated level if previous stimulation occurred, and positive feedback resulting from mutual excitation is the cause. Hysteresis has been shown to occur perceptually in binocular vision.

A second illuminating example of a two-neuron system with sigmoid nonlinearities is one in which the two neurons are mutually inhibitory. If each neuron receives independent stimulation at a different level, strong mutual inhibition will guarantee that the neuron with the stronger stimulus will push the other neuron below its threshold, so only one of the two neurons will remain active. If the stimuli provide evidence for competing hypotheses, this nonlinear network makes decisions. For example, this occurs between competing perceptual representations in the face/vase illusion. Such decision-making networks again require a threshold nonlinearity to produce this result. Decision-making networks can be generalized to any number of mutually inhibitory neurons reflecting any number of competing hypotheses and are referred to as “winner take all” networks because only one neuron wins the competition generated by mutual inhibition. In higher-level cases, the inhibition would occur between competing populations of neurons.

As a third fascinating example of nonlinear dynamics, consider two neurons with negative feedback as described by Equation (5). Thus, the first neuron stimulates the second, and the second inhibits the first. If appropriate values are chosen for k, m, and g in these equations, the response to a constant stimulus will be an oscillation. That is, the R neuron will shift its firing to a high rate in response to the stimulus S, but this will then trigger strong activity in the inhibitory neuron I. This causes the R neuron activity to drop below its threshold, which will then result in the inhibitory neuron dropping below its threshold, so the R neuron can begin firing at a high rate again. This sequence then repeats itself, thus generating an oscillation. These oscillations are known as *limit cycles*. Limit cycles are extremely insensitive to stimulus noise, which makes them ideal to control rhythmic activities. Limit cycles accurately describe perceptual rivalry, such as binocular rivalry or rivalry between interpretations of a Necker cube. Limit cycles also underlie rhythmic motor control activity, including walking, running, swimming, and breathing. Finally, limit cycles can also result if a time delay is introduced into the interaction between nerve cells, and this may be a cause of tremors in certain motor diseases such as Parkinson's. This is an important example of what is called a *dynamical disease*.

In conclusion, perceptual phenomena arise from interactions within neural populations. Single-unit recording only provides incomplete, local clues to population responses but does not explain them. Nonlinear analysis provides a fundamental tool for elucidating how high-level perceptual phenomena may arise from neural population responses. In the absence of nonlinear dynamical models, phenomena such as binocular rivalry, Necker cube reversals, retinal light adaptation, and hysteresis in binocular vision would remain a mystery. As a bonus, nonlinear neural models frequently predict new phenomena, which can then be explored psychophysically and with single unit recording.

Hugh R. Wilson

See also Feedback Pathways; Physiological Approach; Theoretical Approaches

Further Readings

Kaplan, D., & Glass, L. (1995). *Understanding nonlinear dynamics*. New York: Springer.

Trappenberg, T. (2002). *Fundamentals of computational neuroscience*. New York: Oxford University Press.

Wilson, H. R. (1999). *Spikes, decisions, and actions: Dynamical foundations of neuroscience*. Oxford, UK: Oxford University Press.

LOSS OF A SENSE: EFFECT ON OTHERS, PHYSIOLOGICAL

See Cortical Reorganization Following Damage

LOSS OF A SENSE: EFFECT ON OTHERS, PSYCHOLOGICAL

The experience of deaf or blind individuals is so drastically different from what is deemed "normal" experience that, throughout the ages, blind and deaf individuals alike have been treated by societies as having skills that are out of the ordinary—sometimes for the better, but also for the worse. Recent advances in brain science are starting to unveil how blind individuals and deaf people perceive the world differently and what that means for the way they think and interact with their environment.

Generalized Deficiency or Across-Sense Enhancement?

Loss of a sense dramatically alters the type of experience that individuals can rely on as they navigate their world. Some have argued that this is likely to have a negative impact on the sense-deprived individual. Such a view holds that normal development requires integration of information from the different senses, such as when one connects what is being said (audition) with the shape formed by the lips of the speaker (vision). In the absence of such integration, some have argued that development of the remaining senses is compromised and cognitive functioning as a whole is challenged.

Studies of animals and humans show, however, that the loss of one sense is often met by an enhancement of the remaining senses, a phenomenon called *cross-modal plasticity*. For example, cats blinded early in life are better at localizing the source of a sound. They also show enhanced growth of facial whiskers resulting in larger areas of the somatosensory cortex dedicated to tactile exploration. Such compensatory mechanisms often result in deprived individuals outperforming their non-deprived peers on tasks involving their remaining senses. As outlined in this entry, enhanced performance is not systematic. Rather, enhancements in the remaining modalities are most marked under conditions of attention, when the task requires selecting a target from among distractors, or when the time and place of occurrence of the target are unknown. The consequences of these changes for how individuals perceive, remember, and think are being worked out.

The work discussed in this entry only applies to those individuals who have completely lost access to one of their senses early in life, have no associated brain damage, and have been given the opportunity to develop adequate social and communication skills. These selection criteria are necessary to study the effect of sensory loss itself, without being confounded by effects of brain disorder or atypical social development. Thus, most research focuses on blindness caused by early vision loss in both eyes or on deafness caused by profound hearing loss in both cochleas. Although this restricts the relevance of the work to only a small subpopulation of deaf or blind individuals, the focus on early and total deprivation is not arbitrary. It is necessary because brain changes after early deprivation are more widespread. Similarly, brain changes after total sensory loss are more profound and of a different nature than are those observed after partial sensory loss. As research matures, it is encompassing the situation of a greater number of individuals, documenting which mechanisms of brain plasticity are shared by different amounts and onsets of deprivation and which ones may be specific to early and total loss.

Effects of Blindness on Audition, Touch, and Cognition

A review of more than 200 famous blind people from the past few centuries reveals that almost a third of them were professionals or artists in a field related to music. Although such an observation may suggest that blind people have a talent for music, caution is warranted. The prevalence of musicians in the blind population may not result from loss of vision itself. Rather, exposure to music and musical activities may be more likely among blind individuals, indirectly fostering musical ability. This not a mere academic distinction. The view that blind individuals have better musical skills because of their sensory loss holds that sighted individuals will not be able to catch up to the level of blind musicians even if they are given an equivalent amount of music training. In other words, the removal of a sense is necessary for the compensation seen in the remaining senses. On the contrary, the exposure view holds that sighted and blind individuals do not differ in their musical skills and given similar exposure to music, sighted individuals would show the same musical skills as blind individuals. The literature on music skills in the blind is unfortunately rather sparse, leaving this question unanswered. Intriguingly, it appears that absolute pitch, or the ability to identify or produce a note without an external reference, is more likely in musicians who lost their sight within the first few years of life than in sighted musicians or blind musicians who lost their sight later in life.

Better documented is the remarkable compensatory skills that blind individuals display when tested on various auditory and tactile tasks. For example, they are better able to determine where sounds originate from, and they have an easier time distinguishing among complex textures by touch. These changes are more systematically observed in attention-demanding tasks, such as monitoring a stream of sounds for a target sound. Little changes are observed in sensory thresholds, such as the audiogram, which measures the amount of loudness needed to hear different sound frequencies. Enhancements are not restricted to the remaining senses, but are surprisingly widespread. There are reports of blind individuals outperforming sighted counterparts when asked to recall a list of words several days after presentation, or in the study time needed to memorize a long list of words in its exact order of presentation. These enhanced verbal memory skills demonstrate that sensory loss can also lead to cognitive compensatory adjustments.

The performance enhancement in blind individuals goes hand-in-hand with a reorganization of their primary visual cortex (V1). Remember that blindness in these individuals does not result from cortical damage, but rather from loss of functioning of both eyes. As a result, the visual cortex is being denied visual input. Interestingly, recruitment of V1 in blind individuals correlates with performance on tasks as varied as sound localization, tactile discrimination, Braille reading, and verbal memory. In addition, temporary disruption of V1 results in transient impairments of Braille reading in blind individuals, but not in sighted controls. Thus, a brain area typically dedicated to visual analysis has been reorganized to sustain several other abilities as a result of sensory loss. A radical thought emerging from this work is that what was once considered to be the visual cortex may be the cortex that receives inputs from multiple senses. Such multimodal connections are believed to be latent even in area V1 of sighted individuals. Following visual loss, these connections are unmasked, revealing widespread plastic brain changes in this area. The extent of cross-modal plasticity would then be determined by the amount of existing cross-modal interactions at the time of sensory loss. Greater plasticity in early than in late blind individuals is to be expected because redundant connectivity across senses is a landmark of early development. As the individual matures, pruning occurs with only specific neural pathways and local neuronal circuitry being maintained, reducing the multimodal plastic potentialities of the visual cortex by adulthood.

Effects of Deafness on Vision and Language Processing

Several studies report deficient visual skills in deaf individuals. Yet, when confounds are removed and only the effects of auditory loss are studied, a different pattern of results emerges. Deaf individuals do not have better vision than hearing individuals in the sense that they need prescription lenses less often, but for some aspects of vision, such as peripheral visual attention, their skills are heightened. For example, deaf individuals are better than their hearing peers are at localizing a peripheral target in a cluttered visual scene. A real-life example would be the ability to detect a deer standing on the side of the highway while driving. Conversely, when focusing on a central task—such as reading or doing math—their attention is more likely to be pulled away by peripheral events. This behavior has often been described in the literature as increased distractibility and has led many to argue that deaf individuals lack the ability to focus visual attention. Recent studies have established, however, that this distraction effect is not the sign of a deficient focusing of visual attention in the deaf population, but rather of a greater awareness of their surroundings.

Recent evidence also indicates that tactile processing is enhanced in individuals born with profound hearing losses. Again the changes appear to be attentional rather than perceptual. Deaf individuals cannot distinguish between the frequencies of a pair of vibrating rods any better than hearing individuals can, but deaf individuals are better at monitoring a vibrating rod for a change in frequency whose time of onset is unknown. Thus, similar to the plasticity noted in blind individuals, the changes in visual and tactile processing in deaf individuals are also best revealed under conditions of attention.

The neural correlates of these behavioral effects have been linked to a reorganization of the posterior parietal cortex (PPC) and of the posterior superior temporal gyrus (posterior STG). Importantly, these regions are known to be key brain centers for visual attention and zones of convergence for the different senses. Thus, the attentional and multisensory associative cortex reorganizes given missing auditory input. This result is similar to that documented in the animal literature and is well accounted for by a competitive, Hebbian-like mechanism in which neurons that fire at the same time are more likely to maintain their connectivity. When there is no auditory input to these associative areas those neural connections die away, and connections from the remaining modalities take over.

Contrary to the case of blindness, where changes to area V1 have been observed, changes in the primary auditory cortex (A1) have been harder to observe following the loss of audition. It may only be a question of time before such changes are found, or it may be that the connectivity that would allow A1 to reorganize and be taken over by the remaining modalities is lacking. This issue is still being worked out by researchers, but the discrepancy illustrates one of the major challenges in the field of brain plasticity—plastic changes are highly specific in their effects. Although the same neural mechanisms for plasticity appear to be at play across different types of altered experience, the consequence of a given change in experience will be determined by the state of the existing neural connectivity between sensory and associative cortices as well as local neural organization at the time of sensory loss, leading to highly specific patterns of changes as a function of the type and timing of that sensory loss.

Cross-Modal Plasticity, Implants, and Sensory Restitution

The loss of a sense in early childhood is often perceived as a devastating event by parents and peers. Blind individuals are often thought by the sighted to be overwhelmed by darkness. Deaf individuals, because of the language barrier it introduces with the hearing world, have been labeled "deaf and dumb." Even though movies such as *Children of a Lesser God* have raised awareness of what it means to live as a deaf person, deafness is still perceived by many in the hearing world as an extremely debilitating condition. Assuming that individuals who are blind or deaf would welcome becoming sighted or hearing, and leaving cultural issues aside, our understanding of brain plasticity calls for caution. The main determinants of how likely restitutive technologies such as cochlear implants are to be of help are (a) age of onset of sensory loss and (b) age of implantation. Early loss will affect the development of the brain and result in specific cortical reorganization. During development, the brain learns to tune itself to the type of information it receives. The reorganization will be all the more extensive when that loss occurs early and is maintained throughout development. The resulting organization will not be easily undone if the missing modality is suddenly reintroduced in adulthood. By that time, the brain lacks much of the plastic potentiality that would allow it to make use of the added sense. This phenomenon is probably best exemplified by the few reports of early blind and early deaf individuals whose sense has been restored in adulthood. After an initial period of amazement at being able to feel light or sound, these individuals typically report extraordinary difficulties in making sense of the signal provided by their new implant. Many choose to turn off their implant after a few months of struggles.

Implants, however, have had fantastic clinical utility for individuals who suffer from sensory loss in their adulthood. They also have significant restitutive potential for those children who are implanted at a very early age. For example, there is much interest in the conditions that would favor oral language development in deaf children that have received a cochlear implant. An active research effort is ongoing to understand the effects of implantation in babies and toddlers on both the stimulation and recruitment of sensory cortices as well as the proper functioning of higher cognitive processes.

D. Bavelier and Matthew W. G. Dye

See also Attention: Physiological; Audition: Disorders; Cochlear Implants: Technology; Cortical Reorganization Following Damage; Experience-Dependent Plasticity; Multimodal Interactions: Visual–Auditory; Multimodal Interactions: Visual–Haptic; Prostheses: Visual; Visual Disorders: Blindness

Further Readings

Bavelier, D., & Neville, H. J. (2002). Cross-modal plasticity: Where and how? *Nature Reviews Neuroscience, 3,* 443–452.

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences, 10*(6), 278–285.

Kadosh, R. C., & Walsh, V. (2006). Cognitive neuroscience: Rewired or crosswired brains? *Current Biology, 16*(22), R962–963.

Kurson, R. (2007). *Crashing through: A true story of risk, adventure, and the man who dared to see.* New York: Random House.

Rauschecker, J. P. (1995). Compensatory plasticity and sensory substitution in the cerebral cortex. *Trends in Neurosciences, 18,* 36–43.

Turkewitz, G., & Kenny, P. A. (1982). Limitations on input as a basis for neural organization and development. *Developmental Psychobiology, 15*(4), 357–368.

LOW VISION

Low vision refers to deficits in visual function caused by eye disorders. Why should an encyclopedia of perception contain an entry on low vision? Ophthalmologists and optometrists focus on the diagnosis and treatment of eye disease, and on vision care. But vision scientists have recently begun to ask how different types of eye disorders affect visual perception. Because low vision can have a major impact on important daily activities such as reading and driving, interest in low vision has helped propel research on the role of normal vision in these important activities. Low vision also provides informative natural experiments on the effects of visual deprivation. For example, studies using brain-imaging methods have demonstrated that the brain's occipital cortex, normally devoted to visual processing, can participate in tactile perception and other nonvisual functions in cases of severe vision loss. This entry provides a detailed definition of low vision, discusses the demographics, and summarizes the characteristics of the major types of low vision.

Definitions

Low vision can be defined as any chronic visual condition that impairs everyday activities and is not correctable by glasses or contact lenses. In contrast, blindness is defined as the lack of any useful pattern vision.

Four additional terms with subtle distinctions are commonly used: *visual disorder* refers to any deviation from normal in the health status of the eyes, including diseases or chronic conditions. *Visual impairment* refers to consequences of an eye disease or disorder on a clinical test of visual function such as letter acuity. *Visual disability* refers to restrictions on normal activities of daily life resulting from visual impairment, such as reading or mobility. *Visual handicap* refers to a cultural, economic, or social disadvantage for a person resulting from visual impairment.

Visual acuity is often used to define low vision. The World Health Organization (WHO) defines *low vision* as a letter acuity in the better eye worse than 20/60 (metric 6/18), measured with the person's best optical correction (glasses or contact lenses). A person with 20/60 acuity is just able to read letters on the eye chart at 20 feet that a person with normal 20/20 acuity can read at 60 feet. This is equivalent to saying that the observer with low vision would need the letters to be three times larger than the person with normal vision at the same viewing distance.

A definition of low vision based on acuity excludes some people with good central vision but substantial peripheral visual-field loss. For this reason, the definition of low vision usually includes people with major contraction of the visual field.

A disability-based definition of low vision is the inability to read a newspaper at a normal distance (16 inches, 40 centimeters [cm]) with the aid of properly prescribed glasses or contacts. This definition recognizes the importance of reading as a visual activity and the high prevalence of reading problems among people with eye disease.

In developed countries, refractive errors such as myopia (short sightedness), hypermetropia (far sightedness), and astigmatism are not considered to be low vision because they can be routinely corrected with glasses or contact lenses. In developing countries, where optical prescriptions may not be available to many people, these refractive errors can be disabling.

Vision disorders that affect one eye only while leaving the other eye with normal vision, such as amblyopia (informally termed "lazy eye"), do not qualify as low vision.

Low vision should be distinguished from *legal blindness*. Legal blindness—defined in the United States as a corrected visual acuity in the better eye of no more than 20/200, or a visual field of no more than 20 degrees—is used for various legal purposes. The categorization of people as "sighted" or "legally blind" has tended to obscure the many gradations of low vision and the functional value of residual vision.

The term *low vision* was first introduced in 1953 by Eleanor Faye and Gerald Fonda in the United States and has largely replaced earlier terms such as *partial sight* and *subnormal vision*. The term also refers to a specialization in the eye-care professions of ophthalmology and optometry. Low-vision clinics and rehabilitation centers provide services to people with low vision including prescription of magnifiers and other vision aids, training in techniques to accomplish activities of daily life with impaired vision, and sometimes counseling to deal with emotional, social, vocational, and economic issues related to low vision.

Demographics

WHO estimated in 2002 that there were more than 161 million visually impaired people worldwide (more than 2.5% of the world's population), including 124 million with low vision and 37 million who were blind. Contributing to low vision and blindness in developing countries are preventable or curable diseases that are rarely encountered in developed countries. One example is trachoma, a disorder of the lids and cornea spread by flies. Trachoma can be prevented by improved hygiene and treated by antibiotic ointment or simple eyelid surgery. Another example is river blindness (onchocerciasis), also spread by flies but that responds to drug treatment. Even cataract, the clouding of the crystalline lens of the eye, which is routinely treated in developed countries with outpatient surgery, is a major cause of permanent low vision in many developing countries.

The Eye Diseases Prevalence Research Group recently estimated that there are about 3.3 million U.S. citizens older than the age of 40 with impaired vision. Of these, about 937,000 are legally blind and another 2.4 million have milder low vision.

Among the legally blind, approximately 200,000 are totally blind, that is, have no useful pattern vision. Because of demographic trends, particularly the aging of the U.S. population, it is estimated that by the year 2020, the number of people with low vision will increase by 70% to 5.7 million in the United States, with additional people younger than 40 years of age not included in these estimates.

There are age, race, and gender differences in the prevalence of low vision. Because the leading causes of visual impairment in the United States are age-related eye diseases—macular degeneration, glaucoma, diabetic retinopathy, and cataract—the prevalence of impaired vision rises steeply with age. More than two thirds of all people with low vision are older than 65 years of age; 25% of people older than 85 years of age have low vision.

Age-related macular degeneration is the leading cause of low vision among Whites in the United States (> 50%), and cataract and glaucoma among Blacks (> 60%). In the Baltimore Eye Survey, the prevalence of visual impairment and blindness among Blacks was double that of Whites, with the two groups having roughly comparable socioeconomic status.

Although adult men and women are about equally likely to acquire low vision, there are more women than men with low vision because women have a longer life expectancy and prevalence increases with age.

It is difficult to establish an accurate count of the number of people with low vision. Epidemiological factors, such as changes in the age distribution within the population and changes in the incidence of particular disorders, limit precision in counting. For example, in the decade following World War II, many premature infants suffered a blinding retinal disorder called retrolental fibroplasia (RLF) from excessive oxygen exposure in incubators; the blind musician Stevie Wonder is a famous example. As a result, the incidence of low vision and blindness increased. But when the problem was recognized, steps were taken to eliminate this cause of low vision. A similar problem has emerged recently. With the aid of modern technology, infants born prematurely can now survive. Their retinas, however, are immature and delicate, and frequently do not develop normally. The disorder is called retinopathy of prematurity, the modern term for RLF.

Varieties

Low vision is usually discussed in terms of reduced acuity, reduced contrast sensitivity, and visual-field loss. Clinical tests have been developed to measure these “dimensions” of vision. Specific eye diseases vary in their impact on these three dimensions; for instance, cataract often results in reduced acuity and contrast sensitivity, but has little or no impact on the size of the visual field.

Acuity Reduction

Visual acuity refers to the ability of the eye to resolve fine detail. Normal acuity is said to be 20/20 (metric 6/6), although many people with normal vision actually have acuity better than 20/20. An acuity of 20/20 refers to the ability to recognize fine details, such as the gap in the letter C, subtending 1 arc min (1/60 of a degree of visual angle). With suitable magnification, people with acuities as low as 20/2000 (acuity letters 100 times larger than normal) may be able to read slowly. So a range of about 100:1 in spatial resolution covers people from normal acuity to very low vision.

A person’s acuity places a lower bound on the size of objects required for recognition. Here, “size” refers to the visual angle subtended by an object. An object’s visual angle depends on both its physical size and the viewing distance. As an approximation, the visual angle subtended by a target is proportional to the physical extent of the object and inversely proportional to the viewing distance. For example, the width of an adult’s thumb is about 2 cm. Held at arm’s length (about 57 cm), the thumb subtends 2 degrees of visual angle, but at half that distance, it subtends 4 degrees.

A person with reduced acuity will frequently need to enlarge an object’s visual angle to recognize it. Enlarging the visual angle can be done in two ways: by enlarging the physical size of the target, or by reducing the viewing distance. In the case of print, the character size can be enlarged as in large print books, or the reader can bring the text closer to the eye. The latter maneuver frequently requires use of a lens to focus the print at a short viewing distance. A lens used for this purpose is called a magnifier.

Most people with low vision have reduced acuity and benefit from magnification in reading and many other activities. Although magnification is extremely important, it rarely compensates entirely for the effects of low vision. In a study by Gordon Legge and colleagues, reading speed was assessed for 141 people with various types of low vision, with text magnified to more than compensate for their acuity reduction. Despite the ample magnification, only 30% of the participants reached a low-normal reading speed. The conclusion was that visual factors other than acuity reduction were contributing to reading problems.

Reduced Contrast Sensitivity

Contrast sensitivity refers to the ability to see small differences in shades of gray that distinguish one pattern feature from another. A person with good contrast sensitivity may have little trouble reading faded gray letters on a white page whereas a person with reduced contrast sensitivity might fail to see the letters, or might struggle to read them. Research in the second half of the 20th century showed that the visual system specializes in extracting information about the light-dark contrast between objects and their backgrounds.

Many eye disorders, such as advanced cataracts, result in a reduction in contrast sensitivity. Several clinical tests have been developed to measure contrast sensitivity. The Pelli-Robson test is similar to a standard eye chart, except that the letters on successive rows decrease in contrast rather than decrease in size. The further down the chart a person reads, the better is the person's contrast sensitivity.

People with severely reduced contrast sensitivity see only the highest-contrast features in their environment. They fail to see texture on surfaces or shading patterns that help in object recognition. They may miss seeing steps or other low-contrast surface irregularities. Some people with low vision may read slowly, even if the print contrast is high and they have adequate magnification, because of inadequate visual contrast sensitivity.

Contrast polarity refers to the distinction between typical black-on-white text and white-on-black text. For some types of eye disorders, especially those involving light scatter within the optics of the eye such as cataract, people may read better with "reversed contrast" white-on-black text. For this reason, video-based and computer-based electronic magnifiers for low vision usually include a contrast-polarity switch.

Visual Field Loss

The visual field can be portrayed as an island in a sea of blindness. The height of land above sea level represents visual resolution. For the normal eye, the island slopes up to a narrow, steep peak near its center. This is the region of high-acuity central vision. The region of the retina responsible for the high-acuity central vision is the macula, which contains the fovea with its high density of cone photoreceptors. From the center of vision, the normal visual field for each eye extends 60° upward, 75° downward, 100° away from the nose, and 60° toward the nose.

Clinical assessment of the visual field is called *perimetry* and is usually conducted with instruments called *perimeters*. Typically, the patient is asked to watch a fixation target while a test target is presented elsewhere in the visual field. The patient reports on whether the target was seen. The data are compiled into a map of visual-field sensitivity. Modern automated perimeters, such as the Humphrey Analyzer, contain programs to probe different regions of the visual field in an efficient and reproducible manner.

The main reason for clinical measurement of the visual field is to detect and chart the course of visual pathology. Sometimes an abnormal visual field is the first symptom of eye disease, such as glaucoma. Field testing can also be important in understanding the visual capacities of a person with low vision.

Some types of eye disease cause characteristic forms of visual-field loss. Age-related macular degeneration (AMD) is the leading cause of low vision in developed countries. Many patients with AMD suffer such severe damage to the macular region of the retina that they lose all of their central vision. The resulting blind region of the visual field is called a central scotoma. People with this form of low vision must rely on their peripheral vision for seeing. Because acuity is lower in peripheral vision than central vision, people with AMD benefit from magnification. Even with adequate magnification to compensate for their acuity loss, they still read slowly because of intrinsic pattern-processing deficits in peripheral vision.

They may need to avert their gaze to look at objects, including the faces of friends, to image targets in their peripheral field. Fortunately, the remaining function in peripheral vision is often adequate for good mobility.

Other types of eye disease produce a loss of peripheral vision but may spare central vision. Advanced glaucoma and retinitis pigmentosa are diseases that may result in damage to peripheral vision, leaving a narrow central visual field, sometimes termed tunnel vision. People with severe peripheral field loss may retain good acuity and may continue to read without a magnifier. The loss of peripheral vision may produce problems with mobility, both walking and driving.

Other common forms of visual-field loss include patchy visual-field loss in diabetic retinopathy or loss of the left or right half of the visual field (hemianopsia) resulting from stroke or other cortical disorders.

Other Deficits

Other aspects of normal visual perception can be compromised in low vision. Often, vision in one eye is much better than the other. The person is effectively monocular and the advantages of binocularity are lost, including stereoscopic depth perception.

Retinal disease and cataract can result in deficiencies of color vision. These acquired color defects usually differ from the inherited forms of color blindness. In one rare form of low vision, called rod monochromatism, there are no functioning cone photoreceptors in the retina, and affected individuals have no color vision.

Glare from bright lights, including the sun, and difficulties in light adaptation are common in low vision. In retinitis pigmentosa, night vision (scotopic vision) is compromised first in the course of the disease because of damage to the rod photoreceptors in the retina.

Optimizing Low Vision

Until the 1960s, there was a widely held view, termed *sight saving,* that impaired vision should be used sparingly as if it were a nonrenewable resource. Subsequently, sight saving was replaced by the sight-utilization view that even severe low vision has functional benefits, and that these benefits should be optimized. This view has spawned the development of a wide range of adaptive technology for low vision, and the growing awareness of the usefulness of enhancing environmental cues for low vision. It has also led to a growing body of research on the perceptual characteristics of low vision.

Gordon E. Legge and Susana T. L. Chung

See also Ageing and Vision; Contrast Perception; Eye Movements and Reading; Visual Acuity; Visual Disorders: Blindness; Word Recognition

Further Readings

Cheung, S.-H., & Legge, G. E. (2005). Functional and cortical adaptations to central vision loss. *Visual Neuroscience, 22,* 187–201.

Chung, S. T. L., Mansfield, J. S., & Legge, G. E. (1998). Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision Research, 38,* 2949–2962.

Eye Diseases Prevalence Research Group. (2004). Causes and prevalence of visual impairment among adults in the United States, *Archives of Ophthalmology, 122,* 477–485.

Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., et al. (2003). Long-term deprivation affects visual perception and cortex. *Nature Neuroscience, 6,* 915–916.

Goodrich, G. L., & Arditi, A. (2008). Low vision: The reference. Retrieved from http://www.ISLRR.org

Legge, G. E. (2007). *Psychophysics of reading in normal and low vision.* Mahwah, NJ: Lawrence Erlbaum.

Marmor, M. F. (2006). Ophthalmology and art: Simulation of Monet's cataracts and Degas' retinal disease. *Archives of Ophthalmology, 124,* 1764–1769.

Silverstone, B., Lang, M. A., Rosenthal, B., & Faye, E. E. (Eds.). (2000). *Lighthouse handbook on vision impairment and vision rehabilitation.* New York: Oxford University Press.

Vision 2020: The right to sight. Retrieved from http://www.who.int/mip2001/files/1994/Vision2020-TheRighttoSight.pdf

M

MACHINE VISION

See Computer Vision

MAGIC AND PERCEPTION

Most of us have been wowed by magicians and their ability to conjure up rabbits from hats, vanish doves in mid air, or even to saw ladies in half. For decades, magicians have perfected the art of performing acts that seem to defy logic and reality, leaving audiences baffled and amazed. Although these acts appear almost supernatural, they are created entirely by natural means. The performance of magic requires a *method* (how the trick works) to achieve an *effect* (what the spectator sees). Successful magic relies on the spectator experiencing an effect while being unaware of the method. For example, the effect might be the disappearance of a ball. The method involves secretly concealing the ball in the hand while pretending to throw it up in the air. If the magician can produce an effect while preventing the audience from detecting the method, then the audience will experience a magical illusion. Many of the illusions created by the magician are not that different from the types of illusions the mind plays on us in everyday life. The main difference lies in the way in which the inconsistencies between our perception and reality are revealed. People often assume that magicians hide their methods through speedy actions. This idea of the "the hand being quicker than the eye" is far from the truth: Most of the magician's manipulations are carried out at a normal pace. Rather than relying on speed, some of the methods are often concealed through the use of *misdirection* and *illusions,* which are described in this entry.

Misdirection

Misdirection refers to the diversion of attention away from the method so that the audience does not notice how the effect was produced. Indeed, visual scientists have confirmed that we are only aware of a small proportion of the information that enters our eyes. For example, the phenomenon of change blindness dramatically demonstrates that looking at an object does not necessarily mean that we can see it. In order for us to see an object, we also need to attend to it, as most of the information that is not attended to is rapidly lost. Misdirection exploits people's impoverished awareness of the world by systematically manipulating people's attention to prevent them from seeing the method.

There are generally two types of misdirection: physical and psychological misdirection. *Physical misdirection* involves controlling people's attention via stimulus properties. It is now known that certain stimulus properties, such as the appearance of a new object or the sudden physical change of an object,

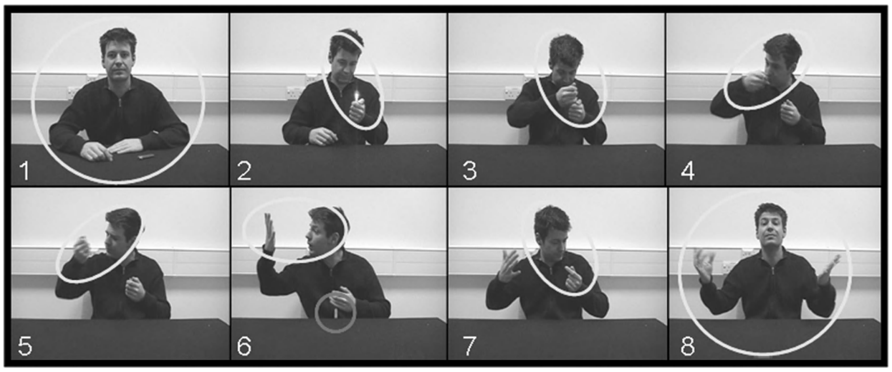


Figure 1 Sequence of a Misdirection Trick

Source: Author.

Notes: The magician picks up the lighter and lights it using his left hand (1–2). He then pretends to take the flame away with his right hand (3–4). Once the right hand has reached the other side, he opens it to reveal that the flame has "disappeared" (5–6). This ensures the spectator's attention has been misdirected from the hand holding the lighter, allowing the magician to drop it into his lap (6–7). Although this is happening in full view, most observers fail to notice the dropping lighter. The effect is a disappearing lighter, while the method involves misdirecting people's attention (white circles) from it being dropped (gray circle).

automatically capture people's attention. In physical misdirection, magicians exploit this automatic attentional capture by developing techniques that allow them to orchestrate people's attention. For example, an important rule in magic states that the audience will look where the magician is looking. Therefore, magicians can use their gaze direction to control where the audience is attending. Magicians may also strategically manipulate stimulus properties that are known to attract people's attention, such as novelty (e.g., making an object appear), movement (e.g., waving a magic wand), high contrast (e.g., sparks and flames), or sound (e.g., finger snapping). Typically, a wide range of techniques are combined to maximize the effectiveness of the misdirection (Figure 1).

Psychological misdirection controls a spectator's attention by manipulating his or her expectations. The aim is to distract the audience from suspecting that a method was being used. A wide range of techniques are generally employed to reduce suspicion. For example, magicians typically keep their audience in suspense as to what they are about to see. As long as the audience does not know what is about to happen, they will be less aware of the important aspects of the routine and thus less likely to attend to the important aspect that involves the method. This is why magicians typically refuse to perform the same trick twice.

Misdirection involves manipulating people's covert rather than overt attention. This is why people may fail to see an event even if they are looking at it. Successful misdirection is indexed by the observer's awareness of the event rather than eye movements. Even though some people may not be misdirected, their resilience toward misdirection would be revealed by the detection of the event, rather than their eye movements. In other words, people who are not successfully misdirected would detect the event. Therefore, it is not really possible to use eye movements as a criterion of whether someone has been misdirected or not.

Illusions

At times, magicians find it useful to distort people's perception of an object or event through the use of illusions. For example, a magician may want you to misjudge the true size of a box in order to prevent you from suspecting that something is being concealed inside. Our visual system involves many intelligent problem-solving processes. For example, to perceive depth, the visual system must reconstruct the third dimension from a two-dimensional retinal image. This reconstruction process is facilitated by applying knowledge that has been acquired in the past. Although the use of this top-down knowledge greatly reduces the computational cost, it can lead to errors, which are typically experienced as perceptual illusions. Magicians often exploit these errors to create some of their illusions. In the earlier example, the illusionist might manipulate the perspectives of a box to distort its true size and thus leaving enough room to hide an elephant.

Sleight of hand magicians tend to rely on "higher level" cognitive illusions, rather than optical illusions described earlier. In these types of illusions, the magician manipulates people's expectations to the

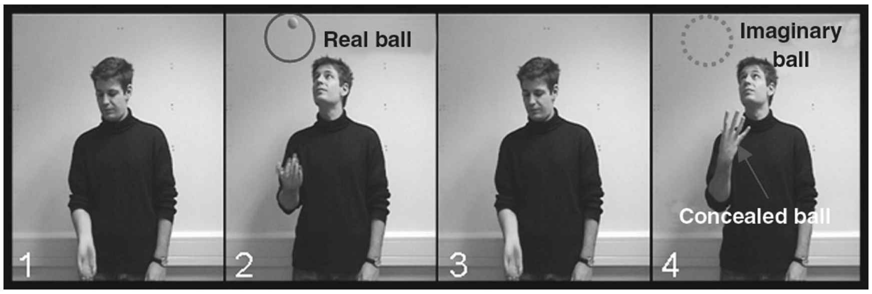


Figure 2 Sequence of the Vanishing Ball Illusion

Source: Adapted from Kuhn & Land, 2006.

Notes: The magician is seen throwing a ball up in the air (1–2). On the final throw (3–4) he merely pretends to throw the ball, while secretly concealing it in his hand. Although the ball does not leave the hand, most of the participants claimed to have seen a ball leaving the hand and move toward the top of the screen. This illusion demonstrates how people's perception is based on expectations, rather than the true sensory input.

extent that they start perceiving what they anticipate instead of what is happening for real. These types of illusions almost take the form of hallucinations. Scientists have recently started investigating such illusions systematically. In one such study, participants viewed a video clip of a magician pretending to throw a ball up in the air. Even though the ball did not leave the magician's hand, several of the participants claimed to have seen an illusory ball move toward the top of the screen (Figure 2). These results demonstrate that their perception of the ball was therefore driven by their expectation, rather than the physical presence of the ball.

Over the centuries, magicians have identified many of our perceptual limitations and have developed techniques that exploit these weaknesses to create their tricks and illusions. The necessary secret nature of magic has meant that most of these techniques remain unknown to the public. Although scientists are starting to understand some of the mechanisms behind these principles, our susceptibility toward these illusions is often outside of our control, thus guaranteeing the survival of this art form.

Gustav Kuhn

See also Attention: Cognitive Influences; Attention: Divided; Attention: Selective; Change Detection; Eye Movements During Cognition and Conversation; Top-Down and Bottom-Up Processing; Visual Illusions

Further Readings

Kuhn, G., Amlani, A. A., & Rensink, R. A. (2008). Towards a science of magic. *Trends in Cognitive Science, 12*(9), 349–354.

Kuhn, G., & Land, M. F. (2006). There's more to magic than meets the eye. *Current Biology, 16*(22), R950–R951.

Kuhn, G., & Tatler, B. W. (2005). Magic and fixation: Now you don't see it, now you do. *Perception, 34*(9), 1155–1161.

Lamont, P., & Wiseman, R. (1999). *Magic in theory.* Seattle, WA: Hermetic Press.

MAGNETOENCEPHALOGRAPHY

How are events in the external world transformed into perceptual experiences via electrical coding in the brain? This simple question forms one of the most basic and long-standing problems of perception. *Magnetoencephalography,* or MEG, is one of several noninvasive brain imaging techniques that allow scientists to explore the link between neural activity and perception.

Like the related technique of electroencephalography (EEG), MEG essentially measures electrical currents generated by neural activity. MEG measures these electrical currents indirectly, through their magnetic fields. (It is a basic principle of physics that moving electrical currents produce magnetic fields.) MEG has excellent temporal resolution, on the order of milliseconds, allowing noninvasive real-time recording of neural activity. Therefore, this technique is well suited to examine the time course of perceptual processing in the brain.

However, in contrast to its high temporal resolution, the spatial localization of MEG is relatively poor. That is, MEG can indicate *when* neural responses occur with great precision, but not exactly *where* the activity takes place. Nevertheless, its millisecond temporal resolution makes MEG a valuable tool for both basic research and clinical applications. In this entry, how MEG works, the differences between MEG and EEG, and MEG measures of perception will be covered.

How MEG Works

As previously described, MEG measures magnetic fields associated with neural activity. These signals are extremely weak: The largest neuromagnetic fields, such as the spontaneous alpha rhythm, are only approximately 10^{-12} tesla (T). Magnetic fields evoked by stimulus presentation are usually even smaller, in the range of tens to hundreds of femtotesla (fT; 10^{-15} T). For comparison, the magnetic field of the earth is about 30 to 50 microtesla (μT; 10^{-6} T), roughly a billion times greater than these evoked neuromagnetic signals. Recording these weak neuromagnetic signals requires highly sensitive superconducting quantum interference device (SQUID) sensors capable of detecting tiny magnetic fields. A single MEG machine typically contains hundreds of SQUIDs arranged in a helmet shape inside a cryogenic dewar, a container filled with liquid helium to maintain the superconductivity of the SQUIDs. Sensors are designed for maximal sensitivity to the nearest source of interest, the brain, excluding magnetic fields located further away. This equipment can also be enclosed in a magnetically shielded room to further exclude ambient magnetic noise from sources such as electrical devices, radiofrequency signals, and the Earth's magnetic field.

Recorded MEG data represents changes in magnetic field strength as a function of time. Perception research often focuses on evoked magnetic fields in response to stimulation, calculated by averaging together tens to hundreds of trials. These evoked responses can be studied either in isolation or in conjunction with source localization. Source localization methods attempt to solve the "inverse problem" of estimating neural sources from the signal recorded at the scalp. Mathematically, the inverse problem has an infinite number of solutions: Any number or combination of neural sources can produce a given signal measured at the scalp. Therefore, MEG cannot achieve the precise spatial localization of neuroimaging techniques, such as functional magnetic resonance imaging. However, additional constraints based on anatomy and known brain functionality, as well as mathematical techniques (e.g., beamforming), can allow better estimates of the origins of MEG signals.

MEG Versus EEG

Although arising from more or less the same neural sources, the signals recorded with MEG and EEG differ in several respects. Compared to electrical currents, magnetic fields are not distorted by intervening bone, skin, and other tissues. Also, due to the perpendicular orientations of the electric and magnetic fields, MEG can only detect signals from sources oriented tangentially to the position of the sensors at the head, anatomically corresponding to sulci. In contrast, EEG records not only tangential signals but also radial signals (perpendicular to the tangential components) arising from the gyri, nearer the scalp. Finally, under the spherical model of brain conductance commonly assumed in MEG and EEG, the magnetic lead field, or magnetometer sensitivity distribution, falls off toward the center of the sphere more quickly than the electric lead field. Thus, MEG is less sensitive to deep neural sources than EEG. Theoretically, these differences should make spatial localization more straightforward in MEG, though this has been debated in real-world testing. Recent research suggests that combined MEG and EEG may yield better characterization and localization of neural sources than either method alone.

MEG Measures of Perception

The millisecond resolution of MEG makes it an ideal method for studying the time course of perceptual processing. Similar to event-related potentials (ERP) in EEG, evoked magnetic fields have been recorded for multiple stages of perception in various sensory modalities. Within the first 50 to 100 milliseconds of stimulus presentation, evoked magnetic fields can be measured from sensors over primary visual, auditory, and somatosensory areas. Later responses are associated with progressively more complex cognitive processes, including visual differentiation between faces and objects, memory-related change detection, attentional monitoring, and language.

MEG has been valuable for understanding the nature of perceptual processing. In the visual domain, for example, research with MEG has shown that perceiving a face proceeds through successive stages of categorization and identification, beginning approximately 100 milliseconds after presentation. Other researchers have combined MEG with neuroimaging to study both the time course and anatomical regions associated with early stages of object recognition. Its high temporal resolution makes MEG particularly suitable for studies of audition and language, such as characterizing the brain processes involved in recognition of words.

Alison M. Harris

See also Brain Imaging; Evoked Potential: Audition; Evoked Potential: Vision; Neural Recording

Further Readings

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmidt, A. M., Dale, A. M., et al. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences USA, 103,* 229–454.

Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics, 65,* 413–497.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience, 5,* 910–916.

Pylkkänen, L., & Marantz, A. (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences, 7,* 187–189.

MAGNITUDE ESTIMATION AND PRODUCTION

See Scaling of Sensory Magnitude

MAPS IN THE CORTEX

See Cortical Organization

MARY THE COLOR SCIENTIST

Mary the color scientist is a version of *the knowledge argument,* a line of reasoning that is widely regarded as a serious challenge to *physicalism* (also known as *materialism*), the popular philosophical theory that the world is entirely physical. Imagine a time far in the future, when the science of color vision has been completed. Everything that can be known through biology, chemistry, physics, and so on has been discovered. Mary is a brilliant scientist who knows all of that information. But she was raised in an entirely black-and-white room and has never seen colors. She learned the science by reading books and watching lectures on a black-and-white television monitor. One day she leaves the room and finally sees colors. Does she learn anything new? It seems intuitively clear that she does: She learns what it is like to see in color. If so, then it is hard to see how the complete physical truth—which she knows before leaving the room—could determine the complete truth about the world. It seems to follow that the world is not entirely physical.

The argument may be divided into two main claims. One says that physical knowledge is not sufficient for knowledge of conscious experience. Call that *the insufficiency claim.* The other says that the insufficiency claim entails the falsity of physicalism. Call that *the metaphysical inference.* A wide variety of considerations have been adduced for and against both claims. This entry describes a representative sample. It also briefly describes the philosophical significance of the debate about Mary.

The Insufficiency Claim

The Mary case was devised to establish the insufficiency claim. The reasoning runs as follows. Mary knows everything physical before she leaves the room. She is a flawless reasoner: She can deduce all logical implications of the science she learns. Therefore, if physical knowledge were sufficient for knowledge of conscious experience, then she would acquire no additional knowledge when she leaves. Yet, intuitively, she *does* acquire additional knowledge when she leaves. The insufficiency claim follows: Physical knowledge is not sufficient for knowledge of conscious experience.

In response, some physicalists argue that the insufficiency claim derives its appeal from underestimating the implications of Mary's physical knowledge. Suppose that, when she leaves the room, she is greeted with a blue lemon. Would she be fooled about what it is like to see yellow? Not necessarily. By using a portable brain scanner, she would be able to tell that her brain processes correspond to those that others outside the room normally have when they see blue. Some argue that the insufficiency claim fails to take such considerations into account. Its defenders tend to respond by granting that Mary is not easily fooled, but denying the intuition that she learns something assumes otherwise.

Other physicalists reject the insufficiency claim by arguing that some physical information cannot be learned by watching black-and-white lectures. If their argument succeeds, then any information Mary acquires when she leaves the room might be physical information. In that case, her increase in knowledge would not indicate that the physical story about the world is incomplete. Physicalism would be safe from the knowledge argument. However, physical information is typically understood to be information discoverable by objective science and expressible in the colorless, mathematical language of physics. There is no clear reason why such information could not be conveyed through black-and-white lectures. The proposed response seems to imply that the physical extends beyond such objective information. And this, some argue, threatens to render the meaning of "physical" so broad as to trivialize the physicalist's claim that everything is physical.

The Metaphysical Inference

The metaphysical inference says that the insufficiency claim entails the falsity of physicalism. The argument for this premise is subtle and complex, but the basic idea is relatively straightforward. If physicalism is true, then the complete physical truth determines all truths, including the truths about what it is like to see in color. So, if physicalism is true, then all such truths should be deducible from the complete physical truth. In other words, if physicalism is true, then physical knowledge would be sufficient for knowledge of conscious experience—which is precisely the claim that the insufficiency claim denies. Therefore, if the insufficiency claim is true, then physicalism is false.

Some physicalists respond by arguing that knowing what it is like consists in possessing abilities, not information—abilities such as the ability to *visualize* red things. On this view, although Mary's knowledge grows when she leaves the room, the growth consists in her acquiring abilities, not information. And that result is consistent with the physicalist's claim that all information is physical. However, many agree that abilities are *part* of what she acquires, but some doubt that abilities are *all* she gains.

Other physicalists reject the metaphysical inference by arguing that Mary's learning consists in acquiring new ways to represent information she possessed before leaving the room. As it is often put, she acquires new *phenomenal concepts* but no new information. This view has been developed in several ways, some of which are intended to undermine the insufficiency claim instead of the metaphysical inference. The main challenge for the view concerns the status of her new concepts: Her conceptual growth must explain her growth in knowledge. The concern is that any concepts that explain her growth in knowledge might incorporate a nonphysical component.

Jackson's Retraction

The classic version of the knowledge argument, for which the Mary case was invented, was developed in 1982 by the Australian philosopher Frank Jackson. In 1998, he surprised the philosophical community by rejecting the knowledge argument and embracing physicalism. But the knowledge argument has no shortage of defenders. Whether or not it ultimately succeeds, the controversy has shed light on a variety of philosophical matters, including the implications of physicalism, the limitations of science, and how consciousness relates to paradigmatically physical phenomena.

Torin Alter

See also Color Deficiency; Consciousness; Inverted Spectrum; Mind and Body; Private Nature of Perceptual Experience; Qualia

Further Readings

Alter, T., & Howell, R. (2009). *A dialogue on consciousness*. New York: Oxford University Press.

Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly, 32,* 127–136.

Ludlow, P., Stoljar, D., & Nagasawa, Y. (Eds.). (2004). *There's something about Mary: Essays on phenomenal consciousness and Frank Jackson's knowledge argument.* Cambridge: MIT Press.

MATERIAL PROPERTIES PERCEPTION

See Surface and Material Properties Perception

MCCOLLOUGH EFFECT

The *McCollough effect* (ME) is a visual aftereffect in which illusory colors are contingent on the orientation of black-and-white lines. For example, after looking at alternating black-and-red vertical lines and black-and-green horizontal lines, black-and-white vertical lines look greenish and black-and-white horizontal lines look reddish (see color insert, Figure 26d).

The ME is fascinating for:

- being simple to establish, yet being long lasting: 10 minutes of induction produces MEs lasting 24 hours;
- being a contingent aftereffect (CAE) in which one property of the visual world, for example, color, is contingent on another such as orientation; and
- offering a key to the binding problem: how different properties of the sensory world, such as color and orientation, analyzed in different parts of the brain, are experienced as bound together into, say, a red vertical line.

This entry covers what it is like to experience the ME, the history of the ME, other contingent effects of ME, and an explanation of the ME.

Experiencing the Effect

One needs some patience to experience the ME. First one looks at a *test* figure, containing regions of black-and-white lines of opposite orientations, such as vertical and horizontal (see color insert, Figure 26a). All the regions should appear black-and-white.

Next, one views *induction* stimuli for about five minutes. One can do this by staring alternately for about 10 seconds each at black lines of one particular orientation on a background of one color (e.g., vertical black-and-red; color insert Figure 26b) and black lines of the opposite orientation on a background of a complementary color (i.e., horizontal black-and-green; color insert Figure 26c).

Finally, when one looks again at the test figure, one should now see that the vertical white parts look greenish and the horizontal white parts look reddish. Similar tints may also be visible in the dark parts. The test figure may now look like Figure 26(d) of the color insert.

History

The first CAE was reported by Ivo Köhler in 1951. After looking through prisms producing bluish and orange fringes on the left and right of vertical contours respectively, he observed illusory complementary colors contingent on whether a vertical edge was left or right facing. That is, left-facing edges appeared orange and right-facing edges appeared blue. Celeste McCollough realized that similar aftereffects could be made contingent on the orientation of contours; in 1965 she reported the effect that came to bear her name.

Other Contingent Effects

After McCollough's paper appeared, numerous other contingent aftereffects were discovered, including: color contingent on size, texture, and motion; orientation contingent on color; and motion contingent on size, depth, and color. For example, after viewing left-moving red stimuli and right-moving green stimuli, stationary red stimuli appear to move right and stationary green stimuli appear to move left. Contingent aftereffects occur in audition (e.g., between direction of motion and change in pitch) and in haptics (e.g., between felt length of a bar and hand). Frank Durgin reviewed the literature in 1996 to show that not all possible pairs of sensory properties can be or have been demonstrated (e.g., color cannot be made contingent on direction of gaze). Since 1996, researchers have added other CAEs, including face aftereffects contingent on weight, orientation, and gender. For example, inducing with fat male faces and thin female faces makes average-weight male faces look thin and average-weight female faces look fat. If the ME is merely one example of a general class of CAEs, there must be a general explanation.

Explanation

McCollough's own explanation involved two main ideas. First, neurons jointly responsive to color and orientation adapt. For example, neurons responsive to red verticals would reduce their activity during induction compared to that of unadapted neurons responsive to green verticals. Second, perception of color is mediated by the ratio of activities of neurons responsive to complementary colors for the same orientation. For example, if the activity of neurons responsive to red verticals exceeds that of neurons responsive to green verticals (as would usually happen if the vertical were red), then red would be perceived; if the activity of neurons responsive to red verticals is equal to that of neurons responsive to green verticals (as would usually happen if the vertical were black and white), then no color would be perceived; and if the activity of neurons responsive to red verticals is less than that of neurons responsive to green verticals (as would usually happen if the vertical were green), then green would be perceived. Because of the adaptation, when both sorts of neurons respond to black-and-white verticals, the reduced activity of red tips the perceptual balance toward green.

Most researchers of other CAEs prefer a similar explanation. For example, neurons (or neuron assemblies) responsive to fat male faces would reduce their activity during induction compared to that of unadapted neurons responsive to thin male faces. When both sorts of neurons respond to average-weight male faces, the reduced activity of fat tips the perceptual balance toward thin. The failure of particular pairings of properties to become contingent could be because there are no neurons jointly sensitive to those properties.

Other general explanations include various types of learning. These appeal because adaptation declines much more rapidly than CAEs, whereas learning endures. But at least one property of MEs seems to rule out learning as it is usually understood: A ME occurs when the induction stimuli are alternated so rapidly that observers see only a yellow plaid. If one color is never consciously associated with one orientation, it does not seem that any association could be learned between them.

Most visual perception scientists believe that CAEs occur from error-correcting mechanisms misled by a concentrated diet of contingent properties. It is as though the visual system assumes that if all vertical contours are red, then there must be some sort of error, such as in focusing of the eyes, that needs to be corrected to restore the usual situation in which colors are independent of orientation. If so, the ME's error-correcting neurons must be low in the visual system, because it is constrained by retinal position, orientation, size, and wavelength, rather than by environmental position, orientation, size, and color. The ME is essentially confined to one eye (if the adapting stimuli are viewed by one eye, no ME can be seen with the other), suggesting the ME's mechanism is before the convergence of information from the two eyes. Some proposed that the errors corrected in the ME are colored fringes from chromatic aberration. In color insert Figure 26(d), real color fringes simulate this correction, producing a good simulation of the experience of the ME: colors spread away from contours to give a general tint.

Robert P. O'Shea

See also Aftereffects; Afterimages; Binding Problem; Color Perception; Eye: Structure and Optics; Face Perception; Haptics; Neural Representation/Coding; Perceptual Learning; Statistical Learning

Further readings

Dong, C.-J., Swindale, N. V., & Cynader, M. S. (1999). A contingent aftereffect in the auditory system. *Nature Neuroscience, 2,* 863–865.

Durgin, F. H. (1996). Visual aftereffects of texture density contingent on color of frame. *Perception & Psychophysics, 58,* 207–223.

McCollough, C. (2000). Do McCollough effects provide evidence for global pattern processing? *Perception & Psychophysics, 62,* 350–362.

Watson, T. L., & Clifford, C. W. G. (2006). Orientation dependence of the orientation-contingent face aftereffect. *Vision Research, 46,* 3422–3429.

MELODY PERCEPTION

Melody consists of a sequence of pitches organized in time. In a melody, it is the pattern of relationships among the pitches that is important, and not the absolute pitch levels. That is, we can pick any pitch on the piano, and if we play a succession of pitches that goes up two semitones, up two semitones, down four semitones, and so on, we will have "Frère Jacques." (To go up one semitone on the piano keyboard, just move one key to the right, including both black and white keys.) We could write the pitch pattern of "Frère Jacques" [+2 +2 −4 0 +2 +2 −4 +4 +1 +2 −3 +1 +2]. The zero indicates where we repeat a pitch. We could start anywhere on the keyboard, and as long as we follow this pattern the result will be "Frère Jacques." This property of being transposable to any pitch level made melody a favorite example of the Gestalt psychologists in the early 1900s to illustrate the idea that the whole was different from the sum of the parts. We can change every pitch in the melody, but as long as the pattern is preserved the melody remains the same. This entry describes the contour, constraints, and perceptual frameworks of melody perception.

Contour

A melody has been described as a dynamic shape in the musical space of pitch and time. The singer can move the melody up or down in pitch, elongate it by slowing it down, condense it by speeding it up, and within broad limits it remains the same melody. Important features of this dynamic shape are its contours of pitch and rhythm—the pattern of relations between successive notes. Beethoven illustrates this at the start of his Fifth Symphony, where the same pitch-and-rhythm contour is repeated at numerous pitch levels and with many different pitch intervals between the two pitches in the pattern. The song "Frère Jacques" in a minor key [+2 +1 −3 0 +2 +1 −3 +3 +2 +2 −4 +2 +2] is still recognizable as "Frère Jacques" (though you can tell that there's something different about it), an effect Mahler uses in his First Symphony. The pattern of ups and downs [+ + − 0 + + − + + + − + +] is important.

Constraints

The human auditory system imposes constraints on melodies. The pitches in melodies must lie within the audible range of frequencies, and, less obviously, must go neither too slow nor too fast. To be easily recognized, familiar melodies must be presented in a range of tempos between about 0.6 notes/second (1,670 milliseconds/note) and 6 notes/second (167 milliseconds/note). Ornamental arabesques and flourishes can go faster (up to about 20 notes/second), but the notes tend to blur into a continuous stream. Phrases in melodies are usually shorter than five or six seconds, fitting easily into our immediate memory buffer for audition.

Other constraints arise from the musical culture. All but perhaps two of the cultures of the world make use of tonal scales—pitch patterns that divide the octave into five, six, or seven pitch categories and establish a hierarchy of pitches in relation to the tonic (in European music, the familiar "*do re mi*" scale). A culture's melodies will use that culture's pitch categories, and it will be difficult for someone brought up with a different scale system to sing them in tune. Melodies generally move by small steps along the scale, only occasionally leaping across pitch categories, and typically begin and end on the tonic pitch. (See the "Frère Jacques" description earlier. Note that European scale steps are usually one or two semitones apart, never more than three.) Hermann von Helmholtz suggested that people need the landmarks of the scale categories to keep their bearings in tonal space—they show us the way back to the tonic "home base."

Perceptual Frameworks

Cultures also have systems for the organization of time: meters that establish a pattern of regularly recurring beats. The perceptual frameworks of the tonal scale for pitch and the metrical organization of time distinguish music perception from perception in other domains, such as the perception of visual shapes or of tastes and smells. They facilitate the speed and accuracy of our perception of patterns that conform to the frameworks. Strongly tonal melodies are better remembered than weakly tonal ones. For example, Leonard Cohen's song "The Sisters of Mercy" is much easier to remember than his "Suzanne." And our judgment of temporal intervals improves markedly when those intervals are presented in relation to a beat.

According to E. Glenn Schellenberg, Paul Iverson, and Margaret McKinnon, people can often identify highly familiar popular songs in less than a quarter of a second. But when they do that, they are clearly not relying on the pitch pattern and the rhythm of the song, but rather on the timbres and texture of the music. If we restrict the cues to just the pitches and rhythm, it generally takes musicians and nonmusicians alike six or seven notes, or about 3.5 seconds. That's still quite fast, considering the very large number of songs that people know.

We often wonder, what makes a melody really memorable? The earlier considerations suggest that conformity to the tonal and metrical frameworks of our particular culture helps. Many highly memorable melodies exemplify that: "La donna è mobile" from *Rigoletto,* Beethoven's "Ode to Joy," Rodgers and Hart's "Where or when," the Beatles' "Hey, Jude." Nevertheless, there must be countless melodies that exemplify those properties and yet rarely come to mind. We still have much to learn about the mind and brain in order to understand what makes some melodies hit our nervous system just right.

W. Jay Dowling

See also Audition: Pitch Perception; Music Cognition and Perception

Further Readings

Dalla Bella, S., Peretz, I., & Aronoff, N. (2003). Time course of melody recognition: A gating paradigm study. *Perception & Psychophysics, 65,* 1019–1028.

Dowling, W. J. (2001). Music perception. In E. B. Goldstein (Ed.), *Handbook of perception* (pp. 469–498). Oxford, UK: Blackwell.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition.* Orlando, FL: Academic Press.

Hannon, E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science, 16,* 48–55.

Jones, M. R., Summerell, L., & Marshburn, E. (1987). Recognizing melodies: A dynamic interpretation. *Quarterly Journal of Experimental Psychology, 39A,* 89–121.

Schellenberg, E. G., Iverson, P., & McKinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin & Review, 6,* 641–646.

Warren, R. M., Gardner, D. A., Brubaker, B. S., & Bashford, J. A., Jr. (1991). Melodic and nonmelodic sequences of tones: Effects of duration on perception. *Music Perception, 8,* 277–290.

Winkler, I., & Cowan, N. (2005). From sensory to long-term memory: Evidence from auditory memory reactivation studies. *Experimental Psychology, 52,* 3–20.

MICROSTIMULATION

Microstimulation is a tool used to study the neural substrates of perception. This entry discusses the procedures, history, mechanisms, and applications of microstimulation. The electrical pulses presented during microstimulation mimic the pulses generated naturally by small groups of neurons. To produce microstimulation, the tip of a fine insulated wire electrode is sharpened, and insulation is removed for several micrometers to expose the metal surface at the tip. The tip of the electrode is positioned within the brain area of interest using both anatomical and physiological landmarks for guidance. Electrical pulses are then generated at the tip of the electrode with the amount of current, frequency of pulses, and duration of the pulse train adjusted to best imitate the activity of neurons in the area stimulated.

The use of electrical stimulation is associated with the very beginnings of the study of the nervous system. In the 18th century, the Italian scientist Luigi Galvani discovered that a charge of static electricity applied to the nerve of a frog's leg produced lifelike twitches. Later, in the 19th century, the German physiologist Emil du Bois-Reymond's use of electrical stimulation of a nerve and its muscle led to his discovery of the action potential—the fundamental unit for electrical propagation along nerve fibers.

Behavioral changes produced by microstimulation can be observed directly, and many of the first uses of microstimulation focused upon overt behavioral changes, rather than potentially covert effects upon sensory perception. In the 1870s, the British neurologist Sir David Ferrier discovered a number of motor fields within the monkey cortex where stimulation produced movements of the eyes, as well as other parts of the body. In addition, he reported that stimulation localized to the monkey's temporal lobe resulted in responses that appeared to indicate the sensation of odors. This may be the earliest demonstration of sensory perception elicited by microstimulation.

More recently, in the mid-20th century, the American neurosurgeon Wilder Penfield used microstimulation to examine the cerebral cortex of human patients who were undergoing surgery to isolate the focus of severe epileptic seizures and developed the well-known sensory and motor homunculi. The sensory homunculus is a description of the allocation of cortical surface to each part of the sensory surface of the body. For example, stimulation of the finger region of the cortical surface caused the patient to report that the fingers were being touched.

A more recent landmark discovery in the use of microstimulation to study perception was the work of Åke Vallbo and colleagues in Sweden, reported in 1984. Performing these experiments on their own bodies, a microelectrode was passed through the skin to make contact with a nerve in the arm. Electrical stimulation parameters were adjusted until the researchers were able to determine that action potentials were being generated in the fiber of a single nerve cell. Most remarkably, the scientist/subject reported a conscious perception of this stimulation that was identical to what could be produced by stimulation of the small receptive field of the nerve on the tip of the finger. This study provided convincing evidence that the signal carried by an individual nerve fiber is propagated throughout the sensory pathway to the cortex and on to the level of conscious perception with high fidelity. This result gave rise to what is known as the labeled line theory of organization in the brain, which says that information is carried in dedicated pathways from the sensory receptor up to the conscious perceptual level.

Moving back to microstimulation of the cerebral cortex, scientists began to report in the early 1990s that microstimulation of sensory cortical areas could have very precise effects upon perception. For example, William Newsome and colleagues in the United States studied an area of the monkey's visual cortex that specializes in processing of visual motion. They used a motion stimulus consisting of large collections of small moving dots of light. For each stimulus, a fraction of the dots moved in the same direction, while the remainder moved in random directions. These stimuli had the special advantage that they could easily be adjusted to change the degree of difficulty for detecting the direction of motion. Under optimal conditions, both monkeys and humans can detect the direction of motion when only 2 to 4% of the dots are moving in the same direction while the remaining dots move in random directions. Microstimulation to a group of cells that preferred specific direction of motion shifted the monkey's perceptual judgment in favor of that direction. This shift amounted to the result that would be obtained by increasing the number of actual dots moving in the same direction by 5 to 20%.

The processes of attention, perception, and action are closely linked. Several recent studies have shown that microstimulation of cortical areas that control movement can have an effect upon attention. In a series of studies of the monkey frontal lobe, it has been shown that weak stimulation insufficient to produce eye movements nevertheless has effects upon areas of the visual cortex that are involved in pattern and color vision. Thus, preparation for a movement focuses attention upon regions of the sensory environment that will be the target of that movement. These findings support what is known as the premotor theory of attention.

Although microstimulation will continue to be a useful tool for studies of perception, one must also be aware of its limitations. The focal patterns of activity generated by microelectrodes have important differences from the patterns of neuron activity generated naturally. Under natural conditions, neuron activity is invoked in entire pathways that begin from the site of sensory transduction in, for example, the tips of the fingers, the nose, or the eyes, and travel through both subcortical and cortical regions of the central nervous system. In contrast, microstimulation injects current pulses at a fixed point along the sensory-motor pathway that can only approximate the natural activity of neurons in that pathway. From the site of stimulation, injected pulses initiate activity further along in the pathway, but also cause the initiation of action potentials that move unnaturally in the reverse direction through the pathway. In addition, microstimulation can only approximate the temporal and spatial patterns of neuron firing that are not fully understood. Perhaps the biggest challenge to the use of microstimulation to further our understanding of perception is rooted in solving the problem of perceptual binding. Although scientists have been able to locate cortical sites involved in processing the individual components of a visual percept, such as pattern, color, and motion, it is unclear how these separate components are combined to form the single perceptual entity that we are aware of at the conscious level. This presents immense technical challenges to the use of focal microstimulation within the complex network of neuron assemblies responsible for perception, and any attempt at revealing a solution to the binding problem is likely to require a combination of techniques including multielectrode microstimulation and recording, functional imaging, and sophisticated behavioral paradigms.

Mark A. Segraves

See also Binding Problem; Cortical Organization; Neural Representation/Coding; Physiological Approach

Further Readings

Cavanaugh, J., Alvarez, B. D., & Wurtz, R. H. (2006). Enhanced performance with brain stimulation: Attentional shift or visual cue? *Journal of Neuroscience, 26*(44), 11347–11358.

Moore, T., & Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature, 421*(6921), 370–373.

Salzman, C. D., Britten, K. H., & Newsome, W. T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. *Nature, 346*(6280), 174–177.

Vallbo, Å. B., Olsson, K. A., Westberg, K. G., & Clark, F. J. (1984). Microstimulation of single tactile afferents from the human hand. Sensory attributes related to unit type and properties of receptive fields. *Brain, 107*(Pt. 3), 727–749.

MIGRAINE

When one hears *migraine,* one tends to think "terrible headache." But a migraine is more than a headache; and in fact, head pain, a common component of migraine, is not a necessary component. A migraine is best thought of as an episodic intense heightening of sensation, especially in those sensory systems served by the cranial nerves. Generally, this results in severe head pain, but heightened sensitivity to light (photophobia), sound (phonophobia), odors (osmophobia), and touch to the skin (allodynia) occur too. This entry discusses the types of migraines, some of their physical symptoms, and migraine research.

Migraines have been described in literature and in medical writings since antiquity. The basic characteristics of migraine headache—its throbbing nature, the fact that it is usually more intense on one side of the head, and its association with nausea and often vomiting along with sensitivity to light, sound, and other sensory stimuli—led to its recognition as a distinct illness centuries ago. Extensive recent epidemiological work has demonstrated what a widespread health problem migraines are. In 2004, the World Health Organization identified the headache as one of the 20 leading causes of chronic disability worldwide.

Although many subtypes of migraine have been identified; however, two account for the preponderance of cases and also represent the forms of most interest in terms of perception. A *migraine without aura* is characterized predominantly by a severe headache meeting the earlier description. A *migraine with aura* involves a similar headache preceded or accompanied by transitory neurological symptoms referred to the aura. The aura typically begins 30 minutes to 1 hour before headache onset and lasts between 6 and 60 minutes. Auras may be experienced either as hallucinations (positive phenomena) or as functional loss (negative phenomena). The most common type of aura is visual (90% of auras). It is important to note, however, that somatosensory auras (experienced as tingling or numbness), olfactory auras (hallucinatory odors), and auras affecting speech, language reception, and other higher order cognitive functions also occur. Surprisingly very few individuals report auras in the auditory (hearing) modality; the reason for this is not at all clear. Migraine aura can also occur alone (*migraine with aura without headache)*. This type of migraine is probably far more common than is recognized because the aura symptoms rarely persist longer than 30 minutes and are not followed by debilitating pain; people often don't report this to their physicians and don't realize that they are experiencing a form of migraine.

A visual migraine aura is a striking perceptual experience. Some auras involve only negative symptoms—the gradual loss of vision across one half of the visual field. So, for example, everything to the right of where one is looking disappears, whereas everything to the left remains intact. However, more common is a positive aura or an aura that combines positive and negative features. Although there are a number of variants of positive aura, the most common one—called a fortification aura because of its resemblance to medieval fortifications—is unmistakable. What begins as a tiny flashing spot of light just off the center of gaze gradually enlarges into a C-shaped arc of zigzag line elements. These lines are typically very bright and in constant roiling motion. Over a period of about 30 minutes, this zigzag pattern gradually spreads out into the periphery on one side, obliterating vision in that region. The hollow of the C is often a region of scotoma or blindness. One is not aware of the vision loss here because the human brain is very good at filling in, completing a missing hole with information "guessed" based on what surrounds it. However, if one takes a pencil and moves its tip into the center of the C, in many cases the tip of the pencil will vanish.

Whereas fortifications are the most dramatic visual auras, auras may also consist of dots of light (phosphenes), visual noise (like television screen static), or "heat waves"—similar to the undulating image one may experience looking into the distance in the desert heat. Visual auras are typically restricted to one half of visual space (indicating that they originate in one hemisphere of the brain), and move out across the field over time. A similar pattern of movement—this time across the body surface—is experienced in somatosensory auras. A feeling of tingling or of numbness gradually creeps up the fingers and then the arm, or across the face. This spatial spread of the percept across continuous regions of skin or of visual space provides strong support for the idea that the underlying cause is cortical spreading depression (CSD)—a wave front of activation followed by profound neural depression—spreading through a sensory part of the brain. The very slow rate of this spread (measured as 3–6 millimeters of brain tissue/minute) suggests that the underlying process involves chemical diffusion (perhaps of potassium), rather than synaptic neural transmission. This dramatic set of events has recently been visualized in the occipital cortex using functional magnetic resonance imaging by a research group at the Massachusetts General Hospital.

Although CSD provides a credible model for the aura phenomenology, the underlying cause of migraines remains elusive. For many years, debate raged back and forth between vascular and neural theories of migraine. Today the trigeminocervical complex (TCC) in the lower brain stem and cervical spinal cord is thought to play a key role. The TCC innervates many structures of the neck and head, including the vasculature of the dura and pia, the protective coverings of the brain. The trigeminal terminals on these blood vessels are known to be an important site of action of the triptans, the most effective class of abortive medications for migraine headache pain, although some drugs of this class also cross the blood-brain barrier and may have central effects as well. The triptans are agonists of the neuromodulator serotonin; serotonin is thought to play a significant role in the migraine both peripherally and in the central pain pathways. A second brain stem region implicated in migraine is the periaqueductal gray (PAG) of the midbrain. Stimulation of and damage to this structure have been reported to induce a migrainelike headache in individuals with no migraine history. The PAG has been proposed to act as a gate on the transmission of painful stimulation through the TCC to other brain regions. Finally, somatic symptoms during migraine headaches (nausea, vomiting) and in the prodromal period leading up to an episode (water retention, excessive yawning) point to the involvement of additional brain stem structures, particularly the hypothalamus. Despite these important clues, the details of this complex neural and vascular network remain to be worked out. Beyond the obvious clinical importance of unraveling the mystery of the migraine, its study offers insight into the intrinsic organization of the sensory pathways (through the phenomenology of auras) and into the routes by which the poorly understood pain system plays a sensitizing role on other sensory processes.

Frances Wilkinson

See also Hallucinations and Altered Perceptions; Pain: Physiological Mechanisms; Vision

Further Readings

Burstein, R., & Jakubowski, M. (2005). Unitary hypothesis for multiple triggers of the pain and strain of migraine. *Journal of Comparative Neurology, 493,* 9–14.

Hadjikhani, N., Sanchez del Rio, M., Wu, O., Schwartz, D., Bakker, D., Fischl, B., et al. (2001). Mechanisms of migraine aura revealed by functional MRI in human visual cortex. *Proceedings of the National Academy of Sciences USA, 98,* 4687–4692.

Sacks, O. (1999). *Migraine* (Rev. ed.). New York: Vintage Books.

Silberstein, S. D., Lipton, R. B., & Dodick, D. W. (2007). *Wolff's headache and other head pain.* Oxford, UK: Oxford University Press.

MIND AND BODY

Taking a bite of a Granny Smith apple releases a torrent of sensory experiences—a tart tang on the tongue, a savory smell in the nose, a crisp crunch at the ear, a firm feel in the hand, and a chartreuse color at the eye. This bite also releases a torrent of neural activity in multiple areas of the brain. It is natural to ask: What is the relation between sensory experiences and neural activity? This question, in various forms, has perplexed philosophers for centuries. It now absorbs scientists in fields such as neuroscience, cognitive science, and computer science. All now agree that there are clear correlations between sensory experiences and neural activity. But, remarkably, there is no consensus on how to interpret these correlations: We have, as yet, no adequate theory of the relation between *mind and body.* This entry will cover the neural correlates of consciousness, philosophical and scientific theories of mind and body, and the scientific work that remains to be done.

Neural Correlates of Consciousness

There are many correlations between sensory experience and neural activity. In the case of vision, for instance, activity in the right hemisphere of the brain is correlated with visual experience in the left visual field, that is, in the field of vision to the left of where one is directly looking. Similarly, activity in the left hemisphere of the brain is correlated with visual experience in the right visual field.

The two hemispheres are normally connected by a large band of neural fibers called the corpus callosum, allowing neural activity to flow back and forth between them. This band has been surgically cut in several patients to treat epilepsy. In an experiment with such patients, one briefly shows a visual image of the phrase *key ring* so that *key* appears in the patient's left visual field and *ring* appears in the right visual field. If patients are asked to say out loud what they saw, they say "ring" and give no evidence of having seen the word *key.* If asked to draw with their left hand what they saw, they draw a picture of a key, and give no evidence of having seen the word *ring.*

This striking result is explained as follows. Speech is correlated with activity in the left hemisphere, whereas control of the left hand is correlated with activity in the right hemisphere. The left hemisphere was shown only the word *key,* and the right hemisphere only the word *ring.* The severed corpus callosum prevented neural activity from being shared between the two hemispheres. This failure to share neural activity is correlated with failure of the sensory experiences correlated with each hemisphere to be integrated into a single sensory experience. This experiment raises the fascinating question of how many selves (i.e., how many subjects of experience) comprise the patient with a severed corpus callosum. If there are two subjects of experience, were there also two before the corpus callosum was severed?

Philosophical Theories

Philosophical theories of the relation between mind and body come in three basic types: physicalist, substance dualist, and idealist. There are many varieties of these basic types.

Physicalism asserts that only physical entities and properties exist. What counts as physical will no doubt evolve as the science of physics evolves. Among the entities and properties currently expected to count as physical are space-time, matter, energy, spin, and charge. Some physicalist accounts of mind and body claim that physical systems, such as the brain, *cause* sensory experiences. Other physicalist accounts claim that sensory experiences *are identical to* brain states or processes. Still others claim that sensory experiences do not exist, and are therefore not caused by or identical to brain processes. Physicalism is the dominant philosophical theory today.

Substance dualism asserts that physical substances and mental substances both equally exist, and that neither can be reduced to the other. Sensory experiences are not identical to brain processes, nor do they depend on the brain for their existence. The relation between body and mind is like the relation between horse and rider. The French philosopher Rene Descartes famously espoused a version of substance dualism in which the mind interacted with the body at the pineal gland.

Idealism asserts that only consciousness and its contents exist. The physical world, including the brain, is among the sensory experiences that are the contents of consciousness. On this view, brain activity is caused by consciousness, and the brain itself exists only as sensory experiences created by consciousness. The British philosopher George Berkeley advocated a version of idealism, summarized by the dictum *esse est percipi* (to be is to be perceived).

Scientific Theories

Contemporary scientific theories of mind and body are almost exclusively physicalist. Most try to pinpoint processes of the brain that are responsible for causing the creation of specific sensory experiences, or for causing the general state of being conscious. Roger Penrose and Stuart Hameroff proposed that quantum properties of special structures within neurons called microtubules are central to the transition from preconscious processing to conscious experience. Francis Crick and Christof Koch proposed that a brain structure called the claustrum is critical to generating consciousness, because of its global connectivity to other parts of the brain. Gerald Edelman and Giulio Tononi proposed that conscious experiences arise from certain patterns of neural activity that loop between the brain's cortex and a subcortical structure called the thalamus.

Serious Scientific Work Ahead

It is widely agreed that the current theories of the relation between brain activity and sensory experience are far from adequate. In each scientific theory, an explanatory gap appears at the critical point where brain activity gives rise to conscious experience. This gap precludes each theory from making empirical predictions that one would expect from an adequate theory. For instance, how, precisely, does a specific pattern of claustrum activation, or of quantum transitions in microtubules, or of looping activations between the thalamus and cortex, give rise to, say, the specific sensory experience of the smell of garlic? How, precisely, must this specific pattern be altered in order to give rise to, say, the sensory experience of the taste of mint? Why, precisely, are these particular patterns required? Why are these patterns accompanied by any conscious experience at all? Why don't they simply go on unconsciously? Tough questions like these constitute the "hard problem" of consciousness. The scientific study of mind and body has progressed to the point where it is ready to face these tough questions. The answers that emerge will surely be intriguing.

Donald D. Hoffman

See also Consciousness; Neuropsychology of Perception; Philosophical Approaches; Vision

Further Readings

Blackmore, S. (2001). *Consciousness: An introduction.* Oxford, UK: Oxford University Press.

Edelman, G., & Tononi, G. (2000). *A universe of consciousness.* New York: Basic Books.

Hoffman, D. (2000). *Visual intelligence.* New York: W. W. Norton.

Koch, C. (2004). *The quest for consciousness.* Englewood, CO: Roberts.

MIRAGES

Mirages (color insert, Figures 27a, b, c) are examples of atmospheric phenomena in which air

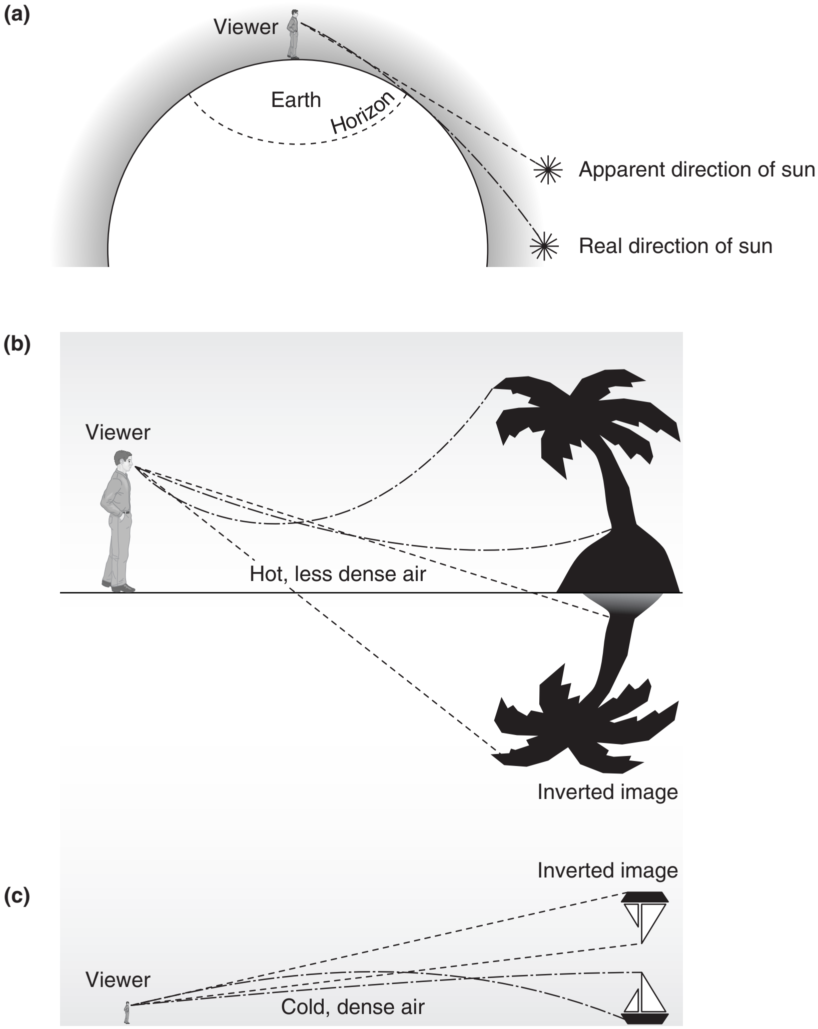


Figure 1 Illustrations of Geometry of Various Refraction Phenomena

Notes: Paths of light are shown as dotted-and-dashed lines. Visual directions are shown as dashed lines. The amount of refraction is greatly magnified. (a) How the sun is geometrically below the horizon when it appears to be touching the horizon. Light from the sun curves down toward the denser air near the earth. (b) An inferior mirage: The normal concave path of light through the atmosphere becomes convex in hotter, less dense air, for example above a desert or roadway. Rays of light that pass through more of this gradient of density are refracted more, essentially yielding reflection of an inverted image below the erect image. (c) A superior mirage: Light from deeper in a gradient of cold, dense air is refracted more than light from higher in that air, essentially yielding reflection of an inverted image above the erect image.

refracts light enough essentially to produce reflections. Other refraction phenomena include the *flattened sun* (color insert, Figure 27d), the *green flash* (color insert, Figure 27e), the *paper-lantern sun* (color insert, Figure 27f), and *twinkling* of distant lights. These are all puzzling phenomena because they are different from what we know about reality: There is no water in the middle of the desert, the sun really is round rather than flattened, its light is white rather than green, and it is continuous rather than irregular or interrupted. Distant lights, such as stars, do not turn on and off. How could these phenomena happen? In this entry, refraction by air and mirages will be covered.

Refraction by Air

Although light travels in straight paths through a constant medium, because the atmosphere is not uniform, light refracts to travel in curved paths. Air is generally denser near the earth's surface; light rays bend toward denser air. One consequence is that when an observer sees the sun just touching the horizon, the geometrical position of the sun is below the horizon (Figure 1a). The apparent direction of an object is determined by the angle that light rays enter the eye, rather than by the object's geometrical position.

If the observer watches the sun carefully as it sets, he or she could see two other refraction phenomena: First, the sun appears flattened (color insert, Figure 27d). Its upper part looks approximately hemicircular, but its lower half is compressed vertically. This is a consequence of the greater refraction of light from the part of the sun lower in the atmosphere; it sends light into the eye at relatively constant angles. Second, when the sun had almost completely disappeared below the horizon, the remaining sliver would appear green, then blue: the green flash (color insert, Figure 27e) and the blue flash. This is because shorter wavelengths, which appear green and blue, are refracted more by the atmosphere; the longer wavelengths, which appear red and yellow, are refracted less, so are blocked from the eye by the curve of the earth.

Both phenomena are more complicated because the atmosphere is never completely uniform. The flattened sun often shows horizontal bands of different lengths (the paper-lantern sun, color insert Figure 27f) and may break into two or more regions of light. The green and blue flashes are very hard to see. These phenomena depend on the properties of the atmosphere that make mirages.

Small distant lights, such as stars and streetlights, sometimes move slightly, appear and disappear irregularly (twinkle or *scintillate*), and change color. This is more pronounced for stars near the horizon than overhead, resulting from inhomogeneities in the atmosphere having different refractive indices. A star's light has to pass through many of these inhomogeneities before it reaches the eye, more when the star is near the horizon. Inhomogeneities change the local direction of the star's light entering the eye (making it appear to move), the amount of light (making it twinkle), and the composition of wavelengths (altering its color). Planets, having much greater visual angles than stars, can send light from different directions into a viewer's eye, making them less subject to these perturbations.

Mirages

When parts of the atmosphere heat differently, such as the hot air immediately above a desert or road, the normal gradient of atmospheric density, the *lapse rate*, with height can be reversed, hot air being less dense than colder air. This reverses the usual concave path of light through the atmosphere: Light entering that hot air at grazing angles can refract so that the outcome is similar to reflection. This gives rise to an *inferior mirage,* called that because it appears below the usual, erect image of the object (Figure 1b). It is inverted, but unlike with true reflection usually involves some parts that are magnified vertically. The classic mirage of thirsty desert travelers is the image of the sky, optically similar to the way a real water surface reflects the sky. The irregular movement of heated air makes the image move and shimmer, appearing like waves on the surface of water, enhancing the illusion.

Mirages can also occur when the air is cooler near the earth's surface, below warmer air, an *inversion layer*. In this case, rays from objects below horizontal bend in the opposite direction (Figure 1c). These *superior mirages* appear above the erect image of the object, although below true horizontal, and are also inverted. Terrestrial objects can include buildings, ships, mountains, or islands; these can be over the horizon.

If the inversion layer is above air having a normal lapse rate, then rays of light can enter that air and bend in the opposite direction. If such rays then enter a viewer's eye, they will form an erect image as if reflected from two mirrors (color insert, Figure 27b). If the rays travel farther before entering an eye, they can re-enter the inversion layer, and then re-enter the other air, and so on. Many alternate inverted and erect images of the object can stack up vertically, giving the appearance of walls and castellations (castle-like images) (color insert, Figure 27c). When these are of islands beyond the sea horizon, they look like a mythical castle: the *Fata Morgana*.

A third type of mirage, proposed by Andrew T. Young and others in 1997, is the *mock mirage*. This is seen when the viewer is above an inversion layer and can appear as alternating inverted and erect images. In the case of the sun near the horizon, these mirages make the sun appear like a paper lantern (color insert, Figure 27f). The uppermost of these mirages can pinch off and magnify the upper rim of the sun, producing pronounced green flashes.

Robert P. O'Shea

See also Air Quality; Atmospheric Phenomena; Constancy; Depth Perception in Pictures/Film; Lightning and Thunder

Further Readings

Lynch, D. K., & Livingston, W. (2001). *Color and light in nature* (2nd ed.). Cambridge, UK: Cambridge University Press.

Minnaert, M. G. J. (1954). *The nature of light and colour in the open air* (H. M. Kremer-Priest, Trans.). New York: Dover.

Young, A. T. (2007). *An introduction to mirages.* Retrieved July 27, 2008, from http://mintaka.sdsu.edu/GF/mirages/mirintro.html

MIRROR NEURONS

Mirror neurons discharge both when a macaque monkey performs an object-directed action, such as grasping a peanut, and when the monkey observes another individual performing the same action, such as when the monkey sees the experimenter grasp a peanut. Thus, these neurons “mirror” others’ actions as if one would perform them oneself, providing a previously unexpected, direct link between individuals. This entry summarizes the most important findings on mirror neurons in monkeys and related findings on mirroring in humans.

Basic Findings

Using single cell recording, mirror neurons have, so far, been discovered in two areas of the macaque brain: first in area F5 of the inferior frontal cortex (ventral premotor cortex) and later in the inferior parietal lobule (IPL). These two brain areas work in concert and receive input from areas involved in the perceptual processing of biological motion, such as the superior temporal sulcus (STS). It is important to note that only a small portion of neurons in area F5 and IPL have mirroring properties.

Two types of mirror neurons can be distinguished, strictly congruent and broadly congruent neurons. Strictly congruent neurons (roughly 1/3 of the mirror neurons in area F5) discharge when the monkey performs a particular action, such as grasping an object with the thumb and index finger (precision grip), and when the monkey observes the same movement. In contrast, broadly congruent neurons (roughly 2/3 of the mirror neurons in area F5) discharge to a wider range of movements during observation. For instance, a broadly congruent neuron may only fire during the performance of a precision grip but not during the performance of a power grip, whereas it may fire regardless of grip type during observation.

It has been suggested that mirror neurons provide a simple and direct form of action understanding. The mirroring mechanism they provide could allow monkeys to understand the actions of their conspecifics by simulating the observed actions using their own action repertoire. The finding that mirror neurons do not discharge when the monkey observes hand actions that are not directed at an object (intransitive actions) illustrates that this understanding seems to be limited to object-directed actions. However, the object an action is directed at does not need to be visible in order for mirror neurons to discharge. If the monkey observes someone grasping an object that is hidden from the view of the monkey, the same mirror neurons will discharge as when the object is not hidden. Thus, knowing that an action is directed at a particular object is sufficient to trigger mirror neuron discharge.

It has been suggested that the two different parts of the mirror circuit, area F5 and IPL, support two different aspects of mirroring. Neurons in IPL seem to be involved with coding the goal of an action, such as eating a fruit or placing it into a container. In contrast, mirror neurons in area F5 seem to code the actual movement that is needed to achieve a particular action goal. Although most studies on mirror neurons have focused on action perception in the visual domain, there is also evidence that some mirror neurons respond to sound. These neurons discharge not only when the monkey sees an action, but also when it hears a sound that normally accompanies a particular action, such as the sound of a peanut being cracked.

Mirroring in Humans

For ethical reasons, it is not permissible to conduct single-cell recordings in healthy humans. Thus, most of our knowledge about mirroring in humans is based on the assumption that brain activation measured with functional magnetic resonance imaging (fMRI) and other techniques can reflect the discharging of mirror neurons.

Numerous brain imaging studies demonstrate that premotor and parietal areas of the human brain are active when people observe others acting. In contrast to the discharge of mirror neurons in monkeys, such activations are not only observed for object-directed actions but also for intransitive actions, such as dancing. The extent to which these brain areas are activated through observation depends on the observer's expertise in performing the observed action. For example, one study investigated whether dancers trained in a particular style would show more activation in brain areas linked to mirroring when observing someone dancing in their own style. Indeed, it was found that ballet dancers showed more activation in the premotor cortex when observing ballet dancing than when observing a martial art style dance that was unfamiliar to them (capoeira). In contrast, capoeira dancers showed more activation when observing capoeira dancing than when observing ballet dancing.

Mirroring in humans has also been implicated in imitation and imitation learning. Several fMRI studies showed higher activation in the human mirror system when an observed action was to be imitated later than when it was observed without the intention to imitate. Therefore, the human mirror system likely contributes to imitation learning. However, it is unlikely that the mirror system is sufficient for imitation: Macaque monkeys do not imitate, although they have mirror neurons. It is likely that imitation learning requires higher-level cognitive functions in addition to the mirror system.

In addition to action understanding and imitation, predicting others' actions has been proposed as another function of mirroring. The match between perceived actions and the observer's own action repertoire may allow the observer to recruit predictive motor mechanisms in order to anticipate what the observed actor is going to do next. For instance, when observing someone throwing a dart at a target board, mirroring may allow the observer to predict where the dart will land.

Future Directions

The wider functional implications of the close perception action link supported by mirror neurons are still under debate. Some researchers have proposed that mirror neurons play an important role in empathy, language evolution, and mental state attribution, leading to the controversial claim that autism is a disorder of the mirror neuron system. In contrast, more traditionally minded researchers hold that mirror neurons are not so special, and that their functionality is restricted to multimodal association. An important contribution to this debate may come from evolutionary biologists who have started to investigate whether other species possess mirroring mechanisms. For instance, new findings indicate that sparrows possess mirror neurons that match perceived songs onto the bird's system for producing songs.

Günther Knoblich and Natalie Sebanz

See also Action and Vision; Animal Motion Perception; Brain Imaging; Embodied Perception; Intentionality and Perception; Neural Recording; Perceptual Development: Imitation; Perceptual Development: Visually Guided Reaching; Reaching and Grasping; Social Perception; Theory of Mind; Visually Guided Actions

Further Readings

Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15,* 1243–1249.

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience, 7,* 942–951.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27,* 169–192.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2006). Mirrors in the mind. *Scientific American, 295*(5), 54–61.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin, 131,* 460–473.

MODALITY (PHILOSOPHY)

Our most basic understanding of the sensory modalities derives from an almost unreflective, casual introspection of our perceptual engagement with the world: Our perceptions of the world seem to come in different categories or classes. These categories—*modalities*—of perceptual experience appear to us as falling into largely distinct groupings and, further, are associated with different bodily organs. For example, when we perceptually interact with a rose, some of our experiences are of a red appearance, a subtle flowery odor, and a soft, velvety texture. Further, these experiences are presented to us in association with our eyes, nose, and fingertips, respectively; if you close your eyes, the experience of the redness of the rose changes, but not the velvety texture. Manipulating the petals of the rose in your fingers is accompanied by more noticeable changes in the appearance of its texture than in its odor. In addition, some take it as given that we find natural continuities and discontinuities among these perceptual experiences that coincide with their groupings with respect to the organs, namely, that the redness of the rose petals and the greenness of the rose stem are closely related to one another (both are *colors*) in a way that neither shares with the smell of the rose.

This extremely basic starting point of this entry begins to generate some of the deeper philosophical questions about the sensory modalities: To what degree are qualities experienced truly different from one another? In virtue of *what* is our sensory experience divided into separate modalities? (These more philosophical questions are different from the more empirical question: To what degree are different sensory experiences related to the activity of specific sensory organs? That is, the former asks a question about the nature of the experiences themselves, whereas the latter asks about the genesis of those experiences.) In exploring questions regarding the nature of the sensory modalities, we are asking questions about the nature of *quality,* in that differences between the senses seem to be intimately tied to fundamental qualitative properties: color, pitch, size, texture, and so on.

Within the Western European tradition, the first systematic account of the nature of the sensory modalities is generally attributed to Aristotle (ca. 384–322 BCE). In his works, *De sensu* and *De anima,* Aristotle presents an account whereby there are five sensory modalities, each uniquely associated with a different bodily organ: sight (the eyes), hearing (the ears), smell (the nose), taste (the tongue), and touch (the skin). Not incidentally, it is probably due to the authority with which Aristotle was granted during the late antiquity and medieval periods that it is now commonplace to refer to any additional sense as a *sixth* sense.

Aristotle's account went beyond what was given to us in brute experience. He proposed that we could differentiate the senses from one another on the basis of those qualities we experience through them. He begins by noting that there are many things that we come to know via our senses, for example, colors, odors, shapes, and sizes. He calls these the *objects* of sense. He then observes that these objects can be divided naturally into different classes. First, consider that some of these objects can be obtained through the activity of more than one sense. Take *shape,* for example; we can both see shape with our eyes and feel shape with our hands. These objects, which also include *movement, rest, number,* and *magnitude,* Aristotle calls the *common objects* of sense, perhaps because they are held in common between multiple senses.

It is a second class of objects that is of most interest here. Aristotle notes that among the objects of sense, there are those that are uniquely and exclusively connected to a single sense. We come to know red and other colors only via the eyes (vision). We come to know bitter and other flavors solely via the tongue (taste). Aristotle calls these singular properties *proper* (or *special*) *objects* of the senses. As Aristotle scholar Richard Sorabji puts it, the proper objects are the things to which the very being of each sense is naturally related. The existence of these proper objects provides a theoretical basis on which to differentiate the senses from one another, in that they provide a principle for the attribution of a sense: If there is some object of knowledge that can only be acquired by the activity of a single organ, then we have grounds for speaking of a specific sense in that case.

Aristotle's proper object account is not without its difficulties. (Aristotle himself grappled with the problem of the proper object of touch—is it temperature or pressure?) Additional accounts of the sensory modalities have been developed in the millennia following the ancient Greeks.

One approach derives from the Aristotelian emphasis on the proper objects of sense. This strategy involves the attempt to differentiate the modalities from one another in terms of what philosophers of mind call *qualia* or what Paul Grice in his chapter "Some Remarks About the

Senses" in the book *Analytical Philosophy* referred to as "the special introspectible character of the experiences [of perception]" (p. 250). This strategy relies on the appeal to the subjective, conscious "feel" of experiences—how the character of the experience of a smell is felt differently from that of the experience of seeing, for example. As Nelson Goodman and Austen Clark describe, one can look to the perceptual science of psychophysics to create "quality spaces" corresponding to the various sensory modalities. Psychophysicists do this by presenting subjects with perceptual stimuli and asking them to identify *just noticeable differences* in what they perceive. This, in turn, allows them to create *matching spans* of stimuli in which each stimulus matches its immediate neighbors, but the ends of the spans are easily discriminable. The quality spaces created by psychophysics have discontinuities that reflect the different modalities: all the colors of the spectrum can be connected together in the same matching span, as can all the pitches in a scale in another matching span; but these two spans are not themselves connected. This discontinuity lends empirical support to the philosophical intuition that the experience of colors is distinct from the experience of pitches, and that vision and audition are separate modalities.

Other accounts eschew the qualia-based approach in favor of a strategy of attempting to use the techniques of neurobiology to identify sensory organs. Instead of beginning with human phenomenology, this approach attempts to discover organs that transmit information to the nervous system, regardless of any conscious correlate to that organ's activity. This allows us to speak of sensory modalities in nonhuman animals, for example, without first answering the question of whether those animals have consciously perceived quality spaces. It also allows for the possibility of discovering putative human sensory modalities that are not conscious, a result that can be taken in two ways: For those that think that the goal of any explanation of sensory modalities is the explanation of our commonsense categories, any account of the senses that allows *unconscious senses* is clearly a nonstarter. For those that believe that the sciences can potentially lead to the revision—even radical revision—of commonsense psychological categories, such neurobiologically inspired conceptions of the senses represent a potentially fruitful endeavor.

Brian L. Keeley

See also Philosophical Approaches; Psychophysics: Detection; Qualia; Theoretical Approaches

Further Readings

Clark, A. (1993). *Sensory qualities.* Oxford, UK: Oxford University Press.

Goodman, N. (1966). *The structure of appearance.* Indianapolis, IN: Bobbs Merrill.

Grice, H. P. (1962). Some remarks about the senses. In R. J. Butler (Ed.), *Analytical philosophy.* Oxford, UK: Oxford University Press.

Grice, H. P. (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Sorabji, R. (1971). Aristotle on demarcating the five senses. *Philosophical Review, 80,* 55–79.

MODULARITY

Modularity implies self-contained, sometimes repeating, units. Modules can be large complex units (such as modules for homes containing large subparts) or small, more uniform units (such as the bricks in the wall of that home). In nature, modular patterns are visible everywhere, from the segmented bodies of colorful caterpillars to the individual toe pads of a dog's foot. Our distant fossil ancestors clearly were modular segmented creatures. We come from a long evolutionary history of animals that have modular body plans. Most of us can agree that the body has different organs, such as the kidney, stomach, heart and lungs, and that each of these organs can be subdivided into modules, such as the four chambers of the heart, but it has been more difficult to parse the brain into modules, particularly that part of the brain that may be most crucial to perception, namely, the cerebral cortex.

This entry explores whether the concept of a module is useful when considering the organization of the mind and brain. One of the initial problems we face concerns the definition of a module. In engineering, a “modular design” implies a system that can be subdivided into smaller parts that can be created and operated independently. In psychology, this term has been variously used to suggest that the mind is made up of specialized mechanisms that evolved to deal with different senses or particular cognitive problems an organism might face. The idea of modular processing also has been used to describe components of basic aspects of perception, such as color and form vision, as well as higher order processes, such as short- and long-term memory. At the level of brain organization, “modular” has been applied to everything from major anatomical subparts, such as the olfactory bulbs concerned with the sense of smell, to specialized cortical areas that deal with complex object perception, to detailed anatomically or functionally defined features, such as cell layers or repeating columns of cells that appear to share common response properties.

Mapping the Mind: A Little History

The issue of whether one can link mental faculties to particular brain modules or locales has a long, detailed, and colorful history. It is beyond the purview of this entry to review this history, but several points are worth highlighting that relate to modern controversies concerning modular localization of function in the brain.

Historically, Francis Gall is closely associated with localization of function in the brain. Gall, born at the end of the 1700s, promoted the idea that the mind is located in the brain and that the mind is modular. He is best known for establishing a system known as phrenology. Phrenologists believed that specific brain modules reflected specific mental traits and that their relative development in the brain deformed the overlying bone, allowing physicians to evaluate the prominence of different faculties by measuring the bumps on the head. This system remained quite popular through the 19th century. When Gall originally argued for brain/mind modules, he was roundly criticized for his heretical views that ran counter to both religious doctrine and scientific proposals about the function of the mind and brain. Although scientists today find little support for Gall’s proposed modular subdivisions of mind, most scientists agree that regions of the brain are specialized for different tasks. Yet, in spite of a large number of studies trying to define the limits of brain and mind modules, there is still no general agreement on their number or exact boundaries.

One can also cite many historical examples of the opposite view, namely, that mind/brain modules do not exist. The work of Karl Lashley represents the clearest historical example arguing against the existence of mind/brain modules. Lashley, born a century after Gall in the late 1800s, used cortical lesions and measured behavior in rats to examine for the locus of memory in the brain. Although Lashley did believe that there were specialized regions of the brain concerned with primary sensation, his conclusion was that the cerebral cortex was equipotential for complex functions, such as memory and learning. In his view, the rats’ memory deficits following cortical lesions simply got worse when lesions were larger, regardless of where the lesions were placed. Lashley’s conclusion was that no brain modules for learning or memory existed and that these functions were based on widely distributed cortical networks. Modern neuroscience recognizes that the brain retains some amount of plasticity well into adulthood and that some functions, such as perceptual learning, can occur in a distributed manner throughout sensory areas.

Modularity in Development and Evolution

Although debate persists about functional modules in mind/brain, current evolutionary and developmental studies strongly support the idea that there are anatomical and physiological modules in the brain. A few examples will suffice to make the point. As previously described, modularity of the body plan appears obvious. Following on the lead of early work in insects, numerous studies in vertebrates have provided evidence that specific sequences of genes control the development of different brain parts. For example, in the vertebrate hindbrain, a set of genes known as the Hox genes are expressed in a sequential fashion during development and their expression matches the development of boundaries of hindbrain compartments (modules). Hox genes are part of an important family of developmental regulator genes that exist in all studied animal phyla and are arranged in clusters in the genome. The forebrain (part of which develops into the cerebral cortex), however, is structurally more cryptic than the hindbrain. Nevertheless, developmental and genetic studies have revealed subdivisions that appear to mark what will later become major areas of the forebrain. The problem with these studies is that they only define large areas of the forebrain and do not neatly subdivide the cerebral cortex into modules that match boundaries of accepted functional areas. Moreover, during development, migrating cells and growing axons (nerve cell processes) freely cross these proposed forebrain developmental boundaries, making it difficult to determine if a true compartmental or modular boundary exists in this part of the brain. Therefore, although most scientists agree that genes regulate the subdivisions of major brain parts, genetically subdividing the cerebral cortex (where complex perceptual and cognitive tasks are accomplished) is problematic.

Current Controversies

As in the case of genetics previously described, most investigators currently agree on many larger brain subdivisions. For example, investigators have known for many years that the cortex contains maps of sensory surfaces. Using microelectrodes, anatomical tools, and lesions, investigators have defined the boundaries of primary cortical areas that map the body surface in the somatosensory cortex in the parietal lobe, the visual world in the primary visual cortex in the occipital lobe, and the frequency of tones in the primary auditory cortex in the temporal lobe. These primary sensory cortical areas exist in all studied mammals. For the olfactory sense, specific areas of the forebrain contain odor maps. Thus, there is general agreement that the cortex can be divided up into large sensory maps or modules. In primates, the importance of these maps to perception is particularly obvious for the primary visual cortex, where a localized lesion causes a blind area (a scotoma) in the corresponding part of the visual world. General consensus also exists that mammals have an orderly cortical map (motor map) devoted to the control of discrete sets of muscles in the frontal lobe.

However, controversy continues concerning boundaries of other cortical areas or modules. The problem is that the cellular architecture of the cortex is surprisingly similar over its entire extent. The cortex consists of ~2 to 4 millimeter (mm) thick layered sheets of cells. Some investigators have argued that the subdivisions or modules of the cortex inherit their regional differences from another area of the brain called the thalamus. The argument here is that cortical modules are defined developmentally by thalamic modules or nuclei. The thalamus has many groups of cells (nuclei) whose architectural boundaries are easy to identify in microscopic tissue sections. These thalamic cell groups communicate with discrete regions of the cortex via long processes called axons. Redirecting developing axons carrying visual information into the part of the cortex that is normally devoted to hearing has allowed animals to “see” without the visual cortex, suggesting that thalamic modules specify cortical areas. However, research also shows that differential gene expression patterns persist in the cortex in genetically manipulated mice that never develop thalamic connections to the cortex, indicating that some architectural modularity is intrinsic to the cortex. How these developmentally defined genetic cortical gradients relate to adult cortical modules is unclear. The other problem is that outside of the primary sensory cortical areas (and a few other regions) there is no general consensus concerning modular boundaries in the adult cortex. Although Korbinian Brodmann, more than a century ago, proposed a map of 52 cortical areas based on cell staining differences, there is still little agreement on the majority of these areas despite massive numbers of anatomical, physiological, and clinical studies.

In addition to controversy over larger regions or areas of the cortex, there is also debate about modules defined on a smaller scale. One could argue that layers define modules because they show distinct cellular connections. There also appear to be anatomically definable modules within layers. For example, in the somatosensory cortex of rodents, the whiskers on the face are represented by discrete clusters of cells called *barrels,* so named because of their shape in cell-stained sections. In the visual cortex of some primates, including humans, axons coming from the thalamus define radial columns of cells that respond preferentially to input from either the left or right eye (ocular dominance columns). Whether these anatomical markers contribute to function remains a controversial issue.

Staining the primary visual cortex for the relative levels of the mitochondrial enzyme cytochrome oxidase shows a distinct pattern of dark dots in the primary visual cortex and dark and light stripes in the secondary visual cortex of monkeys; these dots and stripes have been argued by some investigators to represent functional modules within the primary and secondary visual cortex. Specifically, it has been argued that cells marked by dark and light staining for cytochrome oxidase in the primary visual cortex are specialized for the processing of color and edges, respectively, and that cells that live in these modules, in turn, send their information to cells located in cytochrome oxidase dark and light stripes in the second visual area. The cells within these modules in the second visual area, in turn, put together information about color and form in more complex ways. This information is then transmitted to higher visual areas that contain modules for still more multifaceted processing of object features. Finally, the signals are sent to the inferotemporal cortex where, some have argued, modules exist for particular types of objects, such as faces, houses, or body parts. It also has been proposed that a second hierarchy of visual areas exists concerned with visual space or active vision involving visual motion. Active vision can be considered vision that allows us to guide our body in space or form our hands into the right shape to pick up a cup on the table in front of us. As with the object vision hierarchy (the “what is it” pathway), it has been argued that visual areas related to space (the “where is it” pathway) contain modules that analyze different aspects of motion and space at each hierarchical level. To sum, the broad proposal here is that the two different hierarchies, the different areas within each hierarchy, and the different modules within areas are all specialized to process different pieces of information necessary for the final, goal-directed behavior.

Similar pathways and hierarchies of areas devoted to “what” and “where” processing have also been proposed for the auditory and somatosensory systems as well. Modular hierarchies have even been proposed for frontal cortical areas concerned with decisions about objects and locations.

These proposals remain the subject of active research. One problem with many of these proposals lies in defining anatomical boundaries for the modules and areas within the hierarchies. For most of the cortex, no structural markers that consistently define such areas or modules within areas exist.

Linking Functional and Structural Modules

If we accept, for the moment, the existence of structural brain modules, even if they are hard to define, then we are faced with the larger issue of linking brain modules to mind modules. Although many more experimental tools are available now for linking brain to mind than at the time of Gall, no clear model of this link exists. Some of the more successful efforts to link mind and brain modules have involved electrophysiological recordings in animals trained to perform certain behavioral tasks that specifically engage certain brain areas. For example, it is possible to study the activities of individual cells in the inferotemporal cortex while the animal performs face identification or face discrimination. By relating the activities of individual cells to the decisions made by the animal, it is possible to infer associations between specific brain areas and specific behaviors or cognitive functions.

One limitation of this approach is that so far it has only been possible to study one simple function at a time by recording the activities of cells in one or two different brain areas simultaneously. Therefore, our understanding of how these specialized brain areas and modules work together to facilitate complex perceptions remains sketchy. For example, the visual, auditory, somatosensory, and olfactory signals that arrive at the cortex during a fine dining event result in a holistic culinary experience, yet, very little is known about how information from the different senses come together in the brain to give rise to this coherent experience. Some argue that in order to understand this process, we must solve the so-called binding problem (i.e., figure out how the different modules exchange and integrate the results of the information processed within each). However, whether information from the different modules needs to be actively integrated together in one brain region or through one brain process itself remains an important and outstanding question. The gap in our knowledge about how the different brain modules work together can be attributed in part to the "modularization" of neuroscience itself and to how we conceptualize mental processes.

Even if we can agree on definitions of mind modules and have some idea how they become linked, we still have no clear idea how such mind modules relate to the genetic underpinnings that determine the form of the brain itself. What role, for example, does experience play in determining modules of the mind? It is obvious that the brain is plastic and we learn to attach meaning to perceptual experiences. A person who learns to speak Chinese and not English has limited ability to discriminate between certain English phonemes, a limitation brought about by experience and not genetics. Some investigators have even argued that modules proposed for facial recognition are actually modules designed for discriminating between complex objects that are most commonly encountered in the animal's ecosystem. To wit, an ornithologist may have more cells in this region devoted to discriminating between bird species, while an auto mechanic may have a profusion of cells that code for different models and makes of cars. Clearly, such mind modules must be at least modified, if not defined, by experience. Yet, it is equally clear that mammalian brains are quite similar, and that many mammals can perform very complex perceptual judgments within hours of birth. Wild horses must recognize their mothers and move with the herd within hours of birth. This requires they make enormously complex judgments about forms and motion so they stay with their mother and distinguish her from other moving objects, avoid bumping into others while in motion, or falling over rocks. Clearly, basic modular architecture for complex perception can develop without experience yet must remain flexible enough to allow new information to be stored continuously and later recalled.

Future Directions

Support exists for both brain and mind modules. Brain areas, including the cortex, are not uniform structurally or functionally. Subdivisions can be recognized using a variety of tools, from genetics to functional imaging. Primary sensory and motor areas are the clearest, best studied, and best accepted examples of such modular regional specialization. However, even these subdivisions cannot be considered modules in rigid engineering terms, where modular parts have rigid boundaries and can operate independently. Definitions of brain modules or areas devoted to the representation of more complex cognitive/perceptual properties continue to remain controversial. Much of this controversy stems from our still primitive understanding of the language of the brain. How complex properties are coded by individual or groups of nerve cells still has not been resolved. Appropriately linking mind modules to brain modules will require a better understanding of the neural language. The ability to record from many neurons at once combined with increasing computer power should aid in this endeavor. Additionally, increasing our knowledge of the architectural boundaries defined by genes, protein expression, and connections in a variety of species can guide us as to where to look for functional brain modules. Finally, sophisticated neural models can be used to test the adequacies of hypothesized mind/brain modules.

Knowledge of the functional and structural modularity of the brain has allowed neuroscientists to study each part of the brain in great detail without worrying too much about the complex relationship that one part may bear to other parts of the brain. Analogous relationship between modules also reduces the overall complexity of the problem because results obtained for one module can be extrapolated to others with reasonable integrity. This reductionist approach has yielded valuable details about brain function, but at the same time it has overly "modularized" neuroscience itself. Eventually, conceptual models may need to be developed to explain how the modules functionally cohere into a whole.

Vivien A. Casagrande and
Gopathy Purushothaman

See also Auditory Processing: Central; Cutaneous Perception; Neural Representation/Coding; Physiological Approach; Visual Processing: Extrastriate Cortex

Further Readings

Catania, K. C. (2002). Barrels, stripes and figureprints in the brain—Implications for theories of cortical organization. *Journal of Neurocytology, 31,* 347–358.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology.* Cambridge: MIT Press.

Geary, D. C., & Huffman, K. J. (2002). Brain and cognitive evolution: Forms of modularity and functions of mind. *Psychological Bulletin, 128,* 667–698.

Hubel, D. H., & Wiesel, T. N. (2005). *Brain and visual perception: The story of a 25-year collaboration.* New York: Oxford University Press.

Kaas, J. H. (1982). The segregation of function in the nervous system: Why do the sensory systems have so many subdivisions? *Contributions to Sensory Physiology, 7,* 201–240.

Lashley, K. S. (1950). In search of the Engram. *Society of Experimental Biology Symposium, 4,* 454–482.

Mountcastle, V. B. (1957). Modality and topographic properties of single neurons in cat's somatic sensory cortex. *Journal of Neurophysiology, 20,* 408–434.

Rakic, P. (1988). Specification of cerebral cortical areas. *Science, 241,* 170–176.

Sur, M. (2001). Development and plasticity of cortical areas and networks. *Nature Reviews Neuroscience, 2,* 251–262.

MOLYNEUX'S QUESTION

The celebrated *Molyneux's question* was addressed by the polymath Dublin lawyer, politician, and mathematician William Molyneux to John Locke, and is mentioned by Locke in Book II, Chapter IX of *An Essay Concerning Human Understanding (1690).* The question was as follows:

> Suppose a man born blind, and now adult, and taught by his touch to distinguish between a cube and a sphere of the same metal. . . . Suppose then the cube and sphere placed on a table, and the blind man be made to see: query, whether by his sight, before he touched them he could now distinguish and tell which is the globe, which the cube . . . ? (Hoffman, p. 18)

In considering the question in the light of subsequent research, it is important to be clear that the question is not about the relative role of genes and environment in shaping visual perception. Both Molyneux and Locke assumed that the blind man would have perfectly clear and normal visual perceptions once the optical impediment to vision was removed. Otherwise, the question would be trivial. If the cured blind man could not see the object before him, obviously he would not be able to name it. Locke assumed that what he called the "annexation" of nonarbitrary "primary qualities" and arbitrary "secondary qualities" to optical input was innate. It did not occur to him that we might have to "learn to see" in the modern sense.

The actual question is about what we should now call *visuo-tactile transfer.* Molyneux argued that the newly recovered blind man would clearly see the object before him, but would not be able to name it, because he had names only for objects experienced through touch.

The question became a *cause célèbre* at the close of the 17th century and for much of the 18th, and was debated by philosophers (e.g., Berkeley, Votaire, Condillac, Diderot, and Reid). To see why it involved such a point of principle, it is helpful to jump to George Berkeley's "New Theory of Vision" (1709), where the battleground between the new philosophy and innate ideas is clearly defined. Berkeley begins by asking what would be required of the mind of the person who could recognize by sight an object that he or she had previously only felt. Logically, there would have to be some connection between the representation, or idea (as Berkeley and other empiricist philosophers termed it), of the object in the visual system and in the tactile–haptic system. Basic to the emerging empiricist philosophy was the proposition that sensory impression gave rise to ideas, and that these ideas were the only way in which we could obtain true knowledge of the outside world.

The next stage in the argument is the proposition that there is no resemblance between the ideas derived from different senses, in particular none between the visual idea of an object and its tactile–haptic counterpart. Therefore, we can establish a connection between the two ideas only by experience. Repeated temporal associations between a visual idea and a tactile–haptic idea will cause them to become connected through an "association of ideas," as David Hume was later to suggest in *A Treatise on Human Nature.* The alternative was that the two sensory ideas could be connected through some third idea: a "general idea," transcending the senses. If a general idea transcends the senses, it cannot be derived from experience, and must be innate. Thus, the answer to Molyneux's question must be no, or one would have to concede that there are innate, suprasensory general ideas, such as the "general idea of a triangle," an example which Berkeley caustically derides.

Why were general ideas denounced? The answer is not scientific, but political. Locke was primarily a political philosopher. It has been said, with only little exaggeration, that Locke's "Treatise on Civil Government" was as familiar to the revolutionary American colonists as the Bible. The fundamental cause of the philosophers of the Enlightenment was Liberty, in all its forms. The empiricist manifesto that knowledge comes through the senses should be deconstructed as "use your own experience to find the truth, not what you are told by the ancients." Locke's philosophy was based on opposition to authority, whether from Government or Church. He urged individuals to seek truth using their own intelligence and perception, rather than relying on received opinion.

For the Enlightenment philosophers, general ideas were relics of ancient philosophy. A general idea of a triangle transcending particular sensory representations hinted at Plato and Aristotle, and would be the thin end of the wedge leading to innate ideas of the divine right of kings.

Molyneux's question, then, became a symbol, or sort of rallying cry, for the new philosophy. Blindness in general became a popular theme, often with the slant that the experiences and beliefs of blind people would be quite different from those of the sighted because they had received such different sensory experiences.

There were few actual experiments among this speculation. A report of a cataract operation in 1709 is mentioned by Georges-Louis Buffon in 1749. Later experiments of Roger Grant and William Cheseldson were much discussed. Grant's case was published in the *Tatler* magazine of 1709, and Cheseldon's in the *Philosophical Transactions of the Royal Society* of 1728.

A problem with Molyneux's question in its original form is that complete deprivation of visual experience from birth would require either the absence of a retina, or the complete absence of light. There is evidence from animals reared in complete darkness that the retina does not develop normally, so it is not surprising that such animals show deficits in visually guided behavior. A completely visually deprived animal or person cannot be used to answer Molyneux's question in its strict sense because visual-tactile transfer cannot be tested in the absence of a normally functioning retina. Therefore, modern research has concentrated on individuals with cataracts or other optical impairments which have not prevented light from reaching the retina. As Ione Fine points out, all these cases have had at least some preoperative visual experience. Richard Gregory and Jean Wallace's famous case "SB," for example, is the case history of a man born in 1906 who lost effective sight in both eyes at about ten months of age, and after 50 years as a blind person, received corneal grafts to restore his sight. The consensus from several cases like SB is that the recovered blind can see colors and motion, and even have some capacity for shape recognition. SB, for example, surprised his investigators by being able to tell the time from a large clock on the wall. He could do this, he explained, because he had previously learned to tell the time by touch from a Hunter watch without a glass. On the basis of this observation, the answer to Molyneux's question would be yes because SB demonstrated direct transfer from touch to vision. However, other cases, reviewed by Fine and Michael Morgan, have shown that the recovered blind have large and persistent difficulties in dealing with three-dimensional shape and depth. Even SB seems to have found limited practical use for his vision. In playing darts, for example, he did no better with eyes open than shut.

A recent and direct investigation of Molyneux's question is the work of the Prakash group from MIT working in India under the leadership of Pawan Sinha. Richard Held and three coauthors of that group have recently presented results of their study of three newly sighted young patients who were specifically tested for visual discriminations, tactile discriminations, and for transfer between the two. Their results suggested a complete lack of transfer from normal tactile discrimination to vision immediately after sight onset. Interestingly, however, they did find evidence for touch to vision transfer and of cross-modal recognition, about a week after surgery. The mechanisms of such rapid learning are currently unknown. After over 300 years we still do not have a complete answer to Molyneux's question.

Michael Morgan

See also Multimodal Interactions: Spatial Perception in Touch and Vision; Multimodal Interactions: Visual–Haptic; Recovery of Vision Following Blindness

Further Readings

Chow, K., Reisen, A., & Newall, F. (1957). Degeneration of retinal ganglion cells in infant chimpanzees reared in darkness. *Journal of Comparative Neurology, 107*(1), 27–42.

Gregory, R. L. (1987). Blindness, recovery from. In R. L. Gregory (Ed.), *The Oxford companion to the mind.* Oxford, UK: Oxford University Press.

Gregory, R. L., & Wallace, J. G. (1963). Recovery from early blindness: A case study. In R. L. Gregory & J. G. Wallace (Eds.), *Experimental Psychological Society, Monograph 2.* Cambridge, UK: Heffer and Sons.

Held, R., Ostrovsky, Y., deGelder, B., & Sinha, P. (2008). Revisiting the Molyneux question. *Journal of Vision, 8,* 6,523a. Retrieved from http://www.journalofvision.org/8/6/523/

Hoffman, D. D. (2000). *Visual intelligence: How we create what we see.* NY: W. W. Norton.

Morgan, M. J. (2008). *Molyneux's question.* Cambridge, UK: Cambridge University Press. (Original work published 1977)

MOTION PARALLAX AND STRUCTURE FROM MOTION

As we move about, the spatial relations between our eyes and objects in our environment change in a systematic way. Corresponding changes occur in the spatial relations between the images of these objects that are projected onto our retinas. *Motion parallax* and *structure from motion* describe two ways in which relative motion between the eye and objects and surfaces in the environment can be informative about the shapes of objects and about their layout in the environment. Motion parallax refers to the inverse relation between projected speed and viewing distance, for objects moving at the same three-dimensional (3-D) speed. Structure from motion refers to the changing distances between the projections of objects, or features on objects, that occur when objects rotate relative to the line of sight. Motion parallax and structure from motion are generally considered to be the most important monocular cues to the perception of the shape of 3-D objects and to the layout of objects in 3-D scenes. In this entry, motion parallax, structure from motion, and a combination of the two in everyday perception will be discussed.

Motion Parallax

Motion parallax is based on perspective: When a person moves relative to an object, or an object moves relative to a person, the closer parts of the object will move faster across the retina. If we hold up two pencils at different distances in front of one eye (keeping the other eye closed) and move our head back and forth horizontally, for example, we can easily observe that the closer pencil moves faster relative to the background. Another simple illustration is the motion across the eye of trees at different distances for a person looking out of the side window of an automobile. Because the trees are not actually moving, the variation in the speeds at which the images of the trees move across the retina must be a result of variations in their distances, with the images of the closer trees moving more rapidly across the retina. Hermann von Helmholtz illustrated the powerful effect of observer motion with the example of a person standing in a thick wood who is unable to distinguish from the mass of foliage which branch belongs to which tree or to perceive the separation of the trees. But when the person moves forward, the mass of foliage is disentangled and there is a clear perception of the separate branches and trees, just as in a stereoscopic view. This illustrates the "camouflage-breaking" quality of this depth cue. The use of this cue is as commonplace as moving one's head back and forth slightly while threading a needle. Motion parallax can also occur with a stationary observer and objects moving in the environment.

The information for motion parallax is relative speed in the image projected onto the retina. Objects, or features on objects, that move faster across the retina when the eye is moved are perceived as closer. Using motion parallax to determine the relative distances of objects in the environment is based on the constraint (assumption) that any variations in the projected motions of the objects are due to variations in their distances. Objects, or parts of objects, cannot be moving independently. This is referred to as the *rigidity constraint,* and is central to theoretical accounts of motion parallax. If motion parallax is based on head movements and the objects are stationary, this constraint is likely to be satisfied. But if the objects themselves are in motion, rigidity cannot necessarily be assumed. Two objects moving rigidly do not provide sufficient information for verifying this constraint because any difference in the projected speeds of two objects is consistent with rigid motion. If the number of moving objects is increased and their projected speeds continue to vary in a way that is consistent with rigid motion, it becomes increasingly likely that rigid motion will be perceived. Three objects moving together against a background are almost always perceived as moving rigidly. Motion parallax refers not only to the perception of the distances of separate objects, but to the perception of parts of a single object or features on an object's surface, such as the grain on a wood floor.

When observing the motion of real objects in an everyday environment, we cannot separate the effects of motion parallax from other depth cues, such as relative size and texture. In research designed to demonstrate the effectiveness of motion parallax in isolation, observers were presented with random dot patterns that had no size, texture, or other variations that could be used to determine relative distance. Indeed, when the random dot pattern was stationary relative to the observer, it looked like a flat plane covered with dots. In one condition, the observer moved his or her head back and forth horizontally and dots in the pattern moved horizontally in the same direction or in the opposite direction. The speeds at which the dots moved simulated the motions that would occur if the observer were looking at 3-D surfaces of various shapes. Observers accurately reported both the direction of depth and the amount of depth in the simulated surfaces, demonstrating that 3-D structures can be recovered from motion parallax alone. The results were similar when the monitor moved and the observer's head was stationary, indicating that motion parallax does not require head movements.

Whether we can determine the absolute distance of an object by moving our head back and forth is an unsettled issue. Theoretically, if we know how far our head has moved, the displacement of an object's projection on the retina should provide absolute distance, but it is not clear that human observers can use this information to judge distance.

An especially important instance of motion parallax is *looming.* Looming refers to the change in the retinal size of an object that is approaching the observer. A looming object is almost always seen as approaching. This may be the earliest depth cue to which infants respond, and has obvious implications for survival. Looming is perceived when the projection of an object increases in size in both the horizontal and vertical dimensions. If the increase occurs in only one dimension, size change or rotation is perceived instead. Of course, decreases in both dimensions in the projection of an object will produce the opposite perception, that of recession in depth.

Structure From Motion

The second type of information available in a changing retinal image does not require perspective. It does not depend on variations in distance from an observer, but only on changes in orientation. Consider a pencil oriented horizontally, perpendicular to your line of sight. Now rotate it about its center, keeping it oriented horizontally, until it is pointed straight at you. The distance between the two ends of the pencil in the retinal projection will become shorter as you rotate it. These changes in the projected distances between points undergoing rigid rotation provide information about their relative positions in the 3-D environment. Shimon Ullman coined the expression *structure from motion* to describe this source of information about the 3-D environment available in rotations. It is sometimes called the *kinetic depth effect* (KDE). Unlike motion parallax, this type of information about 3-D structure is available in rigid rotations of objects regardless of the distance of the object from the observer, even if the distance is infinite. Of course, an infinitely distant object would not be visible, but the effects are approximated at large distances, for example, by observing motion through a telephoto lens.

How people use structure from motion to perceive the relative positions of points in the environment has been studied both from a theoretical and an empirical perspective. Theoretically, the 3-D coordinates of as few as four non-coplanar points undergoing a rigid rotation can be recovered from the distances that these points move in as few as three views (snapshots) of the retinal image by solving a set of simultaneous equations. However, empirical studies suggest that people do not use this type of precise mathematical analysis, but use heuristic (rule-of-thumb) processes. An example of a heuristic process found in empirical research on structure from motion is using the difference in the projected speeds of the closest and most distant points observed in motion to determine the depth separation of these points. This process does not recover depth precisely, but may provide adequate approximations in most circumstances.

We know from both mathematical analyses and psychophysical experiments that the shape of objects can be recovered from displays in which projected speed is not affected by distance from the eye (parallel projections), but which part of the object is closer cannot be determined. Consider a transparent glass sphere covered with small dots, rotating about a vertical axis and seen from a distance. In the laboratory, this is simulated using computer animation. Observers easily recognize the shape of the sphere and they can see it rotating about a vertical axis, but if it is far enough away they cannot tell which dots are in front and which are in back. As a result, the perceived direction of rotation is a chance matter. This near–far ambiguity found in structure from motion can be resolved by adding additional cues, such as binocular disparity or occlusion of the back dots by the front dots. Ambiguity in rotation direction was noted in early observations of structure from motion. In the windmill illusion, for example, a distant windmill is observed from an oblique angle. Because it is silhouetted against the sky, the observer cannot determine whether the lowest blade is in front of or behind the millhouse. The direction of rotation of the blades will therefore be ambiguous and may appear to reverse from time to time.

Motion Parallax and Structure From Motion in Everyday Perception

Motion parallax can be isolated in laboratory studies of perception by showing an observer pure translations, without rotation. Structure from motion can be isolated by showing rotations using parallel projections that simulate viewing from a large distance, as through a telephoto lens. But in everyday perception, the two types of information combine to varying degrees. Rotation is almost always present in the everyday perception of 3-D scenes because the direction of gaze typically rotates relative to points in the environment as an observer moves about. Perspective is always present, although it is reduced for distant or small objects. The combination of these two types of information can be illustrated by holding up four fingers on one hand and looking at the fingertips, with the fingertips positioned so that they are not all in the same plane. If we rotate the hand slightly we can observe changes in the speed with which the fingers move relative to the background and in the separations between the fingertips relative to our viewpoint. The closer fingers will move faster in the retinal projection, a motion parallax effect that depends on viewing distance. An imaginary line connecting any two fingertips will change its orientation in 3-D during rotation, and therefore its length in the retinal projection. This is a structure-from-motion effect that is not related to viewing distance. Although the distinction between motion parallax and structure from motion is important for understanding the information available to the observer, our perceptual processes appear to be attuned to using a combination of these two sources of information for the perception of depth from motion.

Myron L. Braunstein

See also Action and Vision; Depth Perception in Pictures/Film; Motion Perception; Pictorial Depiction and Perception; Self-Motion Perception; Virtual Reality: Vision; Visually Guided Actions

Further Readings

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century-Crofts.

Braunstein, M. L. (1976). *Depth perception through motion.* New York: Academic Press.

Domini, F., Vuong, Q. C., & Caudek, C. (2002). Temporal integration in structure from motion. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 816–838.

Ni, R., Braunstein, M. L., & Andersen, G. J. (2005). Distance perception from motion parallax and ground contact. *Visual Cognition, 12,* 1235–1254.

Rogers, B. J., & Graham, M. E. (1979). Motion parallax as an independent cue for depth perception. *Perception, 8,* 125–134

Todd, J. T., & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics, 48,* 419–430.

Ullman, S. (1979). *The interpretation of visual motion.* Cambridge: MIT Press.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology, 45,* 205–217.

MOTION PERCEPTION

The detection and recognition of motion is one of the key functions of vision. This is not a statement only about human vision: In every species where evidence for visual motion processing has been sought, it has been found, from mammals such as monkeys, cats, and rabbits, through amphibians, crustaceans, and insects. Bees and flies, for instance, have been shown to have exquisitely engineered motion processing systems for flight control and for identifying the structure of their surroundings. This entry will discuss the functions of visual motion processing, motion perception as a distinct submodality, elementary motion detectors, sampled motion, eye-head motion, encoding speed and direction, global motion, segmentation and transparency, and integration of direction and the aperture problem.

Functions of Visual Motion Processing

Movement within the retinal image may arise either from the motion of objects around us or from changes in viewpoint as our eyes move around. Information in the pattern of visual movement from these two sources serves at least eight distinct and important perceptual functions.

1. ***Trajectory judgment.*** We can analyze the three-dimensional (3-D) path of an object—a vehicle or a volleyball—and determine whether it will pass through a given location and collide with other stationary or moving objects and what actions we should initiate to intercept or avoid it.

2. ***Segmentation and grouping.*** A very basic and pervasive problem in perception is to determine what elements of the visual array belong together as parts of a single object or surface, and which arise from different objects and so should be processed separately. One of the most effective indicators is motion. Elements that share common movement are perceived to belong together, whereas a discontinuity in motion between two textured regions gives a powerful sense of a boundary between different surfaces, often at different depths.

3. ***Structure from motion*** and ***motion parallax.*** When an observer moves, surfaces at different distances produce different speeds of movement in the visual array (motion parallax). For example, as we look out of a car window, nearby objects sweep past faster than more distant ones. Similar effects occur when viewing a moving or rotating 3-D object. These effects patterns of visual motion can give a strong indication of 3-D structure, the relative depth of different objects, and the solid shape of individual objects.

4. ***Registration of self-motion.*** When an observer is moving through the environment, the large-scale pattern of optic flow provides information about that motion. For an observer who is looking in the direction they are moving, all points in the field of view have a streaming motion outward from a stationary focus of expansion. This is the point toward which the observer is heading, and the rate of expansion indicates when they will collide with the surface on which it lies. If the observer is looking in another direction, or rotating as well as moving forward, the pattern is more complex, but the direction of heading and time-to-contact can still be recovered. Although the optic flow pattern depends jointly on the 3-D structure of the environment and the observer's own movement, both 3-D structure and self-motion can be separately extracted from the pattern (Figure 1a).

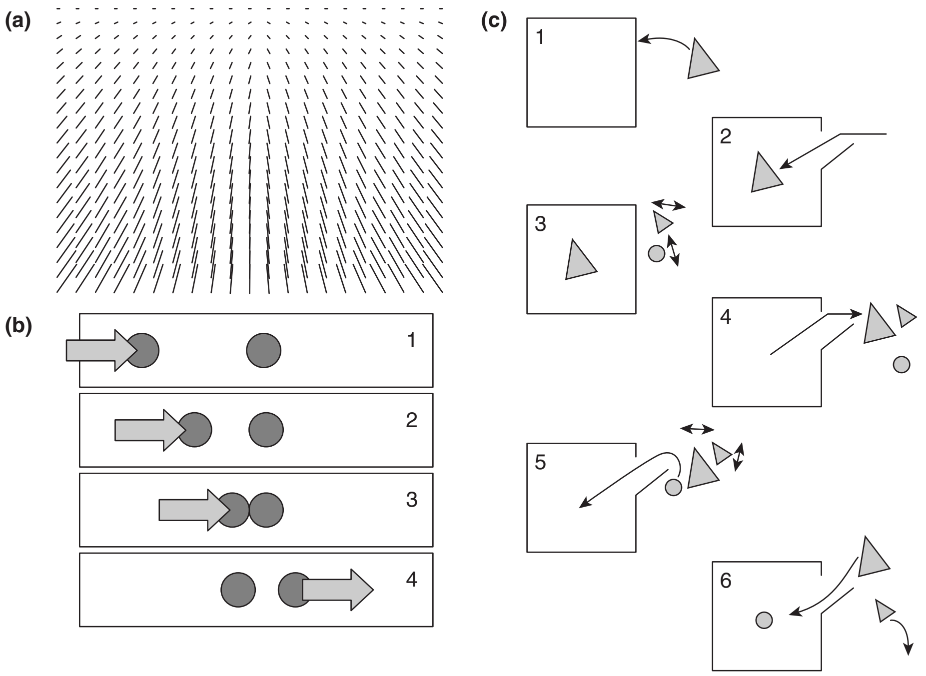


Figure 1 Motion Patterns That Yield Complex Information

Notes: (a) The optic flow pattern generated by moving forward through an environment with a horizontal ground plane, and a vertical obstacle right of center. The short lines represent the direction and speed of image motion at each point. This pattern allows an observer (i) to judge his or her direction of forward motion; (ii) to appreciate that the ground plane is flat with its most distant points near the top of the image area; (iii) to separate the obstacle from the background by the discontinuity in motion at its edges; and (iv) to judge that the obstacle is near and that its front face is flat. (b) Four frames from a motion sequence demonstrating perception of a causal launching event; the left-hand object appears to collide with the right-hand object and set it into motion. (c) Six frames from a motion sequence that gives perception of a social interaction between the participating shapes (Heider & Simmel).

The optic flow arising from self-motion does not simply provide information; it can feed the control of actions related to this self-motion. For a person standing upright, any bodily sway moves the eyes, generating optic flow that can trigger postural correction. Steering and braking, whether driving a vehicle or walking, can be controlled using the pattern of expansion, and the derived time-to-contact.

5. ***Control of eye movements.*** If the whole field of view moves uniformly, this is probably due to rotation of the eyes, either within the head or along with head or body movement. A special-purpose neural mechanism detects such movement and triggers corrective eye movements (the optokinetic response), which acts to stabilize the visual world on the retina. This is one of the most pervasive mechanisms throughout the evolution of the visual system. However, if one object of interest is moving within the field of view, we need to stabilize its image on the retina by moving the eyes, even though the rest of the field will then be filled with contrary retinal flow. This engages a different eye-movement mechanism, using target motion to control a smooth pursuit movement that keeps the target stable on the retina.

6. ***Biological motion.*** Human beings and other animals walking or performing other actions (dancing, throwing, bending, swimming, etc.) produce characteristic patterns of biological motion. Experiments pioneered by Gunnar Johannson reduce the image to a point-light display, with a light attached to each joint or extremity but otherwise filmed in darkness so that no other information is provided on the shape of the person. Such experiments show that observers are extremely sensitive to these movement patterns. They can not only identify the action, but can also frequently use the motion to discriminate the person's gender, individual identity, and emotional state.

7. ***Event perception.*** Biological motion provides specific examples of classifying a type of event from the pattern of motion—a dynamic analogue of how we use static geometry to classify the shapes of objects. Such perceptual classification from motion is not limited to bodily actions. For example, work by Albert Michotte showed that specific relations in the timing of two shapes' movements led to perception of a causal event, in which a colliding object appeared to initiate the second object's motion (Figure 1b). Social interactions (pursuing, accompanying, confronting, evading, etc.), as well as physical interactions, can be inferred from motion, as shown by research following Fritz Heider and Mary-Ann Simmel using simple moving shapes that contain no static cue to social or human identity (Figure 1c). Many other object and event categories, animate or inanimate, have their own characteristic, identifiable motion signatures (consider falling, bouncing, or a machine such as a photocopier operating).

8. ***Motion texture.*** In pattern vision, local features may be combined either to define a global shape (cylinder, slab, etc.) or to yield a perceived texture (e.g., fallen leaves, curly hair) in which the exact position and relationship of the features need not be registered. Similarly, categories of motion can be perceived and encoded, which depend on local motion sequences, but for which their exact layout and timing are irrelevant. Examples would be splashing rain, activity of ants around a nest, or popping kernels in a popcorn dispenser. Such spatiotemporal motion textures have been little studied, but clearly have a vivid perceptual reality in conveying the appearance of particular activities.

Motion Perception: A Distinct Visual Submodality

Mathematically, motion is the change of position over time. Therefore, it might be argued that motion perception derives from the perceived position of visual features. However, this account is inadequate. Clear evidence comes from the motion aftereffect: Prolonged viewing of consistent motion (e.g., a waterfall or a continuously rotating spiral) leads to a vivid sense of movement in the opposite direction when the motion stops. This perceived motion is not accompanied by any corresponding change in perceived position, and so presumably results from adaptation of a neural system that encodes the direction and speed of motion, separately from the systems on which we base positional judgments. However, perceived position and motion are not wholly independent. If part of a display contains a fixed window, within which there are moving stripes, the position of the window appears displaced in the direction of motion. That is, movement determines perceived position as well as position determining motion.

The idea that motion can be considered a distinct visual submodality is linked to neurophysiological findings of a specialized network encoding visual motion. A subset of neurons in the visual cortex (notably, but not only, in the area first identified by Semir Zeki, and known alternatively as medial temporal, MT, or V5) show selective responses to motion in specific directions. Such neurons are predominantly in a network of areas within the dorsal stream of cortical visual processing—a system believed to be engaged in representing the spatial layout of the environment and guiding spatially directed actions. However, activity related to biological motion perception is seen in areas less closely linked to the dorsal stream, and may represent a bridge between dorsal motion processing areas and ventral stream areas involved in identifying people and their expressions.

Elementary Motion Detector: First- and Second-Order Motion

Elementary motion detectors (EMDs) respond to the direction of motion within a small region of the visual field. Various models of human EMDs have been proposed, but they have in common a sensitivity to *motion energy*—a measure of how far the pattern of light in the image shows a systematic shift over time in a particular direction. This measure assumes that the relevant aspect of the moving image is its pattern of light and dark. However, it is possible to perceive motion when this pattern shows no systematic movement. For example, imagine an animal moving through long grass that can only be seen from the moving disturbance in the grass. The average light intensity remains the same over the whole grass, and so there is no motion energy, but yet a clear sense of something moving. This is an example of second-order motion. It appears that motion can be registered from many different attributes of visual stimuli—for example, brightness, contrast, texture, color, or flicker—and that although the initial motion detection process may be different, these processes converge to give a common, cue invariant sense of motion.

Sampled Motion

Real moving objects move continuously, but we can get a convincing sense of motion from an intermittent sequence in which objects appear in a succession of stationary locations. The "moving" images we see on cinema, TV, and computer-graphics screens are examples. Many names have been associated with this phenomenon (apparent motion, beta-motion, phi-phenomenon), but *sampled* or *stroboscopic* motion are neutral and accurate terms that describe the presentation of successive samples along a continuous movement path.

EMDs will respond to sampled motion within certain limits on the gaps in space and time between successive samples. Therefore, the maximum gaps tolerated for perceived motion in patterns of dots have been taken to indicate the spatial and temporal structure of human EMDs. However, when an isolated, identifiable object appears in two successive locations, motion can be seen with much larger gaps in space and time. This led to the proposal of a *short-range process* in motion perception corresponding to activation of EMDs, and a *long-range process* requiring higher level processing that identifies a moving target. In recent work, this distinction appears less clear-cut—certainly the maximum spatial gap can vary greatly depending on the pattern. However, two conclusions seem to be established: (1) a strong distinction remains, especially in terms of time intervals—random dot motion perception breaks down for gaps longer than 100 milliseconds (ms), while isolated targets are seen vividly in motion with gaps at least 2 to 3 times as long; (2) the latter process involves visual attention, and the perceived motion may follow tracking of the focus of attention. A third conclusion is that while second-order motion may be distinct from activation of EMDs, in some cases it involves analogous mechanisms (albeit with different detailed properties), whereas for other attributes (e.g., motion of regions defined by pure color difference or by stereoscopic depth) the attentional tracking process has to be invoked.

Eye-Head Motion

The preceding sections suggest that a common perception of motion can arise from several distinct processes. A further example comes from eye movements pursuing a target. These generate an *efference copy,* which determines that the background appears stationary despite its movement on the retina; at the same time the target, stationary on the retina, is seen to move. Head movements have a similar effect. Human motion perception must be able to seamlessly unify these and many other different sources of information.

Encoding Speed and Direction

As with other aspects of vision, individual neural motion detectors are believed to each respond to a range of speeds and directions. However, we can make quite precise judgments of speed and direction, suggesting that our perception depends on a *population* or *distributed code* using the pattern of activation across many detectors with overlapping sensitivity ranges. This view of direction coding is supported by motion adaptation effects: Prolonged viewing of one motion direction distorts the apparent direction of other motions, making them appear to diverge away from the adapted direction. This is just what would be expected if the pattern of a distributed code was distorted by reduced sensitivity of the adapted detectors.

However, effects for perceived speed are not quite parallel. Rather than shifting perception away from the adapted value, adaptation to motion in the same direction generally reduces apparent speed, whether the adapting speed is slower or faster. The encoding of speed appears to depend on comparing activity in a small number of channels, perhaps just one "slow" and one "fast."

As well as effects of adaptation to previous stimuli, motion perception is also affected by other motions simultaneously present in the field of view. In part, these may be understood as interactions between nearby motion detectors, but a broader principle seems also to be at work. Motion is judged relative to a surrounding frame

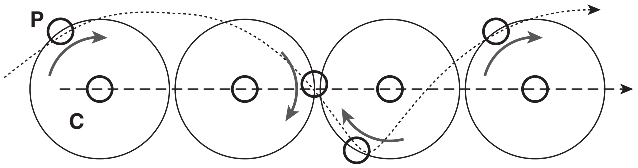


Figure 2 Four Successive Positions of a Rolling Wheel, With Lights on the Perimeter (P) and the Center (C)

Notes: Light P follows a "bouncing" trajectory through space. However, when seen together with the motion of the center light C, its rotary motion is perceived.

of reference—as when the moon, surrounded by wind-blown clouds, appears to be moving in the sky. This phenomenon is known as *induced motion.* Frame-of-reference effects can be more complex, as in Figure 2, which shows the trajectory of lights attached to the center and perimeter of a rolling wheel. Although the light on the perimeter has a bouncing path in space, the context of its shared motion with the center light leads to a compelling perception of its rotary motion in a moving circle. Point-light displays of biological motion, mentioned in an earlier section, provide more complex examples of individual motions represented within a structural hierarchy of motions.

Global Motion: Coherence Sensitivity

Elementary motion detectors each respond to the direction and speed of motion within a small region of the field of view, corresponding to the small receptive fields of neurons in the primary visual cortex (area V1). However, this local information must be combined over larger areas. This is important first because individual EMDs are subject to random noise, and much more precise information about motion can be obtained by combining their results. Second, many of the uses of motion information require integration of information over space to analyze extended patterns of motion (e.g., the optic flow patterns produced by self-motion).

The integrative or global nature of motion processing can be measured using *motion coherence thresholds.* A pattern of dots, all moving with the same direction and speed, is perceived as a uniformly moving surface. If, say, half these dots are replaced with randomly moving dots, the motion is no longer uniform, but there is still a strong and accurate impression of its dominant direction. As the percentage of randomly moving dots is increased, a threshold can be found at which an observer can still just correctly judge leftward versus rightward movement. This coherence threshold is remarkably low—typically when only 5 to 15% of the dots are moving in the same direction. At these low percentages, individual local motions are quite uninformative, and the judgment must depend on combining many local motions to find the statistically predominant direction. There is evidence that such combination occurs in single neurons in cortical area V5/MT (and perhaps other cortical areas also).

Coherence thresholds have proved a powerful tool for analyzing motion processing systems. For example, they can be measured not only for motion in a single direction, but also for the expanding and rotating patterns of optic flow generated by self-motion. The measurements show that optic flow detection integrates motion information over very large regions; these threshold measurements are consistent with cells in the medial superior temporal (MST) area, whose receptive fields are so large that they can cover more than half of the field of view. Coherence thresholds also allow the same performance measure to be used for different motion tasks. For instance, the coherence necessary for children to detect a motion boundary develops differently from that needed to elicit optokinetic eye movements, suggesting that two different neural systems underlie these types of motion-based behavior.

Coherence thresholds have also proved to be a revealing indicator of abnormality, particularly in developmental disorders. Motion coherence can be compared with an analogous task of identifying patterns of short static lines, when a proportion of the lines are given random orientations. Dyslexia, Williams syndrome, autism, and fragile X syndrome all show specific deficits in global motion coherence that go beyond any effects seen in the static analogue. It appears that the visual motion system, or at least its integrative function, is particularly vulnerable to neurological disorder in development.

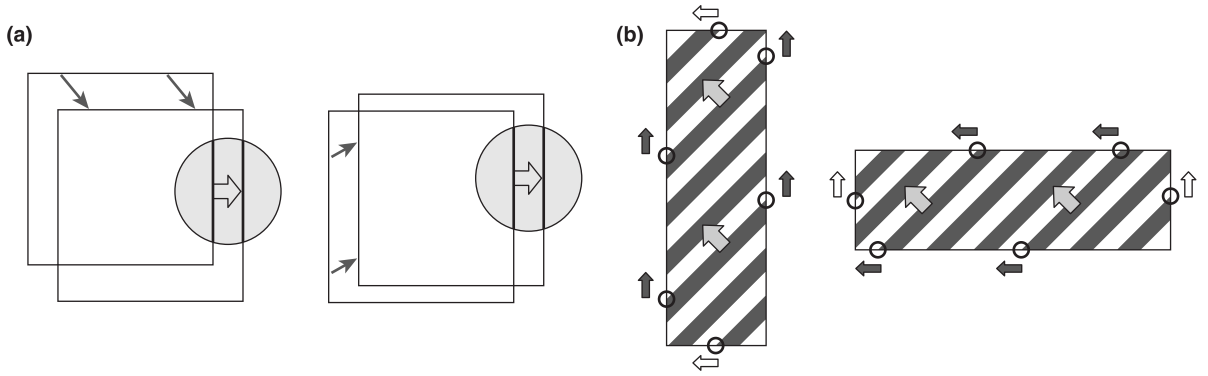


Figure 3 Integration of Motion Directions

Notes: (a) The aperture problem. Two different oblique motions of a square figure are presented. The shaded regions represent the receptive field of a local motion detector. Because this receives information only from the vertical contour, it would respond to the same horizontal motion in each case. Information from other receptive fields, which include the horizontal contours, must be integrated if the true directions are to be registered. (b) The barber pole illusion. Diagonal motion of stripes within a rectangular window is seen as vertical (left figure) or horizontal (right figure). This results from the dominant direction of motion of the circled terminators, where the stripes meet the edge of the window. (Shaded arrows indicate the true motion of the stripes, black arrows show terminators moving in the dominant direction, and white arrows show terminators moving in the perpendicular direction.)

Segmentation and Transparency

The ability to detect motion in low-coherence displays depends on a high degree of spatial integration that combines local motion signals. On the other hand, when dots in adjoining regions move in different speeds or directions, a sharp boundary is perceived, implying that nearby local motions can be processed separately and contrasted. The motion system must therefore be highly flexible, able to combine local motions or to keep them separate, depending on the nature of the stimulus and the demands of a specific visual task.

However, there are limits to this flexibility. If the regions containing distinct motions are too small—for instance, if dots moving left are randomly intermixed with an equal number moving right—the ability to see separate regions breaks down. However, integrative processes do not then simply average all the motions, with leftward and rightward movements canceling out. Rather, there is a perception of transparency, with opposite motions belonging to two visible surfaces, sliding over each other. This implies that the visual system cannot simply represent motions by a single vector at each point in the field of view. There must be a more complex representation that can include multiple motions, each associated with an object or surface that may overlap partly or completely.

Integration of Direction and the Aperture Problem

The requirement to integrate local motions is not simply a matter of improving precision by summing local signals of a common direction. Rather, local motions are intrinsically ambiguous because of the aperture problem illustrated for a diagonally moving square in Figure 3(a). EMDs with localized receptive fields will detect only the horizontal motion of a vertical edge or the vertical motion of a horizontal edge. Both of these are compatible with the overall motion of the square, but only by combining them can the correct direction be unambiguously determined. Like the spatial integration discussed in the previous two sections, in the monkey's brain this integration appears to take place by combining signals from V1 neurons in the specialist motion area V5/MT. This integration is quite a complex process, as illustrated by the barber pole illusion of Figure 3(b). Here, the direction of movement of the bars is resolved by the movement of the terminators where each bar joins the edge of the window; the motion shared by the greater numbers of terminators (e.g., vertical motion when the long edges are vertical) dominates the perceived direction. However, this depends on the bars being perceived to move behind the window; if the bars appear to be in front, the terminators no longer determine the perceived motion. Thus, the system computing visual motion, while specialized, is not wholly self-contained—perceived motion (and that registered by V5/MT neurons) is determined by the perception of depth and shape.

Overall, this is further evidence that motion perception involves a complex and specialized system, receiving and integrating information through diverse routes, and passing it on for a wide variety of perceptual functions.

Oliver Braddick

See also Action and Vision; Aftereffects; Animal Motion Perception; Corollary Discharge; Event Perception; Modularity; Motion Perception: Physiological; Motion Perception: Social; Neural Representation/Coding; Selective Adaptation; Self-Motion Perception; Visual Processing: Extrastriate Cortex

Further Readings

Albright, T. D., & Stoner, G. R. (1995). Visual motion perception. *Proceedings of the National Academy of Sciences USA, 92,* 2433–2440.

Borst, A., & Egelhaaf, M. (1989). Principles of visual motion detection. *Trends in Neuroscience, 12,* 297–306.

Braddick, O. J. (1980). Low- and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 290,* 137–151.

Braddick, O. (1993). Segmentation versus integration in visual motion processing. *Trends in Neuroscience, 16,* 263–268.

Braddick, O., Atkinson, J., & Wattam-Bell, J. (2003). Normal and anomalous development of visual motion processing: Motion coherence and ‘dorsal stream vulnerability.’ *Neuropsychologia, 13,* 1769–1784.

Bruce, V., Green, P. R., & Georgeson, M. A. (1996). *Visual perception. Physiology, psychology & ecology* (3rd ed., chap. 8, pp. 192–204). New York: Psychology Press.

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision, 4,* 103–129.

Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews. Neuroscience, 4,* 179–192.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics, 14,* 201–211.

Lu, Z. L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 18,* 2331–2370.

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1986). The analysis of moving visual patterns. In C. Chagas, R. Gattass, & C. Gross (Eds.), *Pattern recognition mechanisms* (Experimental Brain Research Series, Suppl. 11, pp. 117–151). New York: Springer.

Nakayama, K. (1985). Biological image motion processing: A review. *Vision Research, 25,* 625–660.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4,* 299–309.

Smith, A. T., & Snowden, R. J. (Eds.). (1994). *Visual detection of motion.* London: Academic Press.

Snowden, R. J., & Freeman, T. C. (2004). The visual perception of motion. *Current Biology, 14,* R828–R831.

MOTION PERCEPTION: PHYSIOLOGICAL

Motion is a sine qua non of life. Because of this property, the visual detection of motion is of great significance, particularly for animals. Potential predators, prey, and mates all move, and, insofar as visual systems can detect this at a distance, they will be of great benefit to their owners. Thus, motion automatically captures our attention and generates orienting responses. Moreover, the motion of things is an excellent predictor of where they will be in the future, thus allowing animals to anticipate trajectories and program appropriate movements to either acquire or avoid the moving object. However, because there are many spurious sources of motion—such as those generated by eye, head, or body movements of the observer—distinguishing “real,” object-based motion from self-generated motion is a difficult problem. Visual systems have mechanisms to distinguish these sources of motion and to make use of the different kinds of information they convey. In this entry, basic motion computation, the integration of correspondence and aperture problems, self-motion versus object motion, and links to perception will be discussed.

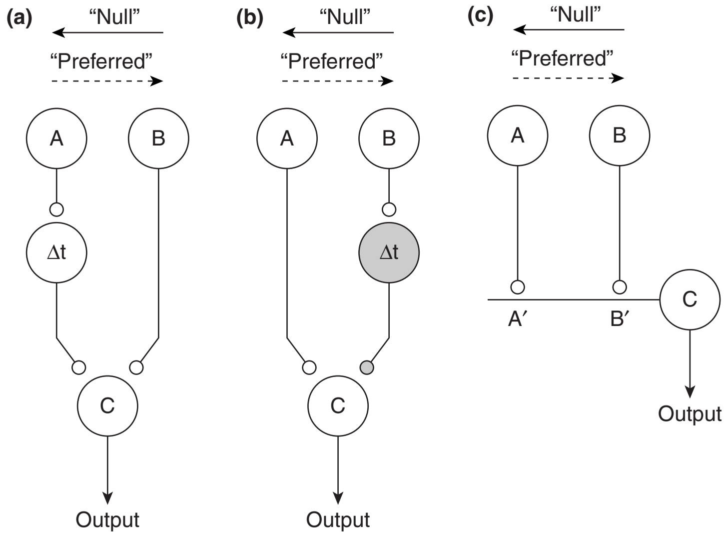


Figure 1 Three Elementary Motion Detectors

Notes: For all three, the direction of motion that maximally activates the output ("C") is from left to right, designated as the "preferred" direction. The opposite, or "null," direction produces no response. Excitatory elements are white; inhibitory elements, gray.

Basic Motion Computation

Motion is displacement over time and thus involves a comparison of spatial location at two different time points. An elementary motion detector (EMD) can be implemented biologically, using neurons, in a variety of ways. To get an idea of how this might be done, consider a circuit that will signal motion in a particular direction, say, from left to right (Figure 1). One needs at least two spatially offset elements that can signal when an "object" is at a given location (to keep it simple, suppose this object is just a spot of light) and some asymmetry in the way that they interact. Neurons signal with electrical impulses consisting of brief perturbations in the potential difference across their cell membranes. So, a spot of light moving from left to right (Figure 1a) will activate neuron A some time before it will activate neuron B. If we now introduce a time delay (Δt, in the form of an extra neuron in the chain) for the impulse originating from A, we can arrange it so that an A-before-B sequence will cause the impulses from A and B to arrive nearly simultaneously and thus sum (or better, multiply) to produce a large response in neuron C. If the reverse, B-before-A, sequence occurs, the delayed impulse from neuron A will now arrive at C well after that from neuron B, and no summation (multiplication) will occur. The same effect can be achieved by using delayed inhibition from the opposite side (Figure 1b) or by having the two signals from A and B arrive at different points on neuron C's asymmetric input process (Figure 1c).

There are many elaborations on these basic types of direction-sensing circuits. For example, one can produce more robust direction selectivity by taking two of the EMDs shown in Figure 1(a)—one of which has the excitatory delay flipped to neuron B's side and so prefers motion from right to left—and subtracting them from each other. Another approach, which appears to be used in some brain regions, is to combine two different mechanisms, such as the asymmetric dendrite (Figure 1c) and the delayed inhibition (Figure 1b), into a single circuit. Recent evidence suggests that these circuits are not genetically hard wired into the brain, but rather require visual experience to develop properly.

Integration: Correspondence and Aperture Problems

Whichever mechanism is used to calculate the direction of movement, there is an obvious problem. The elementary circuits depicted in Figure 1 respond to activation of neurons A and B in a particular temporal order, but there is no guarantee that the two activations were produced by the same visual feature. In other words, how can the brain know that the sequential activations produced at A and B correspond to one moving thing and not two different things, one stimulating A and the other stimulating B? (This is known as the *correspondence problem.*) In a certain basic sense, the answer is that it cannot, and this is the basis of the "trick" used to simulate motion on television and computer screens: None of the pixels on the screen actually move—motion is simulated by turning pixels on and off in a particular sequence, producing what is known as *apparent motion.* Because our EMDs are fooled by certain sequences of pixel flashes, we think we see continuous motion where none is present.

One important version of the correspondence problem concerns recovering the true direction of motion of moving edges. Figure 2 depicts a solid line (broken up into individual pixels for illustrative purposes) moving up and to the right as captured by two successive snapshots at times t_0 and t_1. If we are constrained to sample this motion through an aperture that obscures the bar's endpoints, our EMD has no way of knowing whether it should match the truly corresponding pair of black pixels, the black with the gray, or some other combination. (Remember that the pixels cannot be distinguished in the real situation.) This matters because the different matches yield radically different measurements of the motion of the edge. And this is a real problem for physiological motion circuits because they do, in fact, sample the visual world over a very limited range: Any given direction-selective neuron responds to motion over a small area of the retina, known as its *receptive field.* In the previous example, the EMDs covering the center of the bar can only measure upward motion, so, obviously, the responses from more than one EMD must be combined in some clever way to get the right answer.

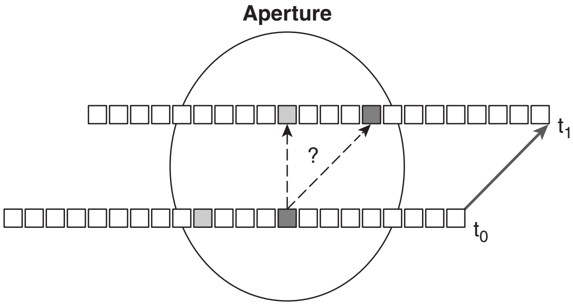


Figure 2 The Aperture Problem for Visual Motion

It is not clear exactly how the visual system solves this problem, but we know that higher order motion processing neurons (i.e., neurons that can pool inputs from many thousands of EMDs) in a region of the monkey's brain called the medial temporal (MT) area do solve the problem. In some models, this is achieved by pooling the responses of EMDs tuned to particular combinations of orientation, direction, and speed; in others, it is done by selectively integrating signals from motion detectors that only signal the direction of discontinuities—such as the ends of the bar—and effectively ignore the middle regions. Given the catholic nature of the brain's bag of neural circuits, it is probable that both mechanisms (and likely others as yet undiscovered) are combined to yield the most robust solution.

Self-Motion Versus Object Motion

Another problem for the EMDs is that rightward motion of an object with respect to an observer can be created in two fundamentally different ways: either the object can move to the right while the observer remains stationary (*object motion*), or the observer can move to the left while the object remains stationary (*self-motion*). Because the EMDs are fixed to the retina, they are subject to multiple sources of observer motion brought about by eye movements (with respect to the head), head movements (with respect to the body), and movements of the entire body.

There are two basic ways that the brain distinguishes self-motion from object motion. One is based purely on comparisons made within the image, and the other makes use of knowledge about the movement made. When retinal motion is

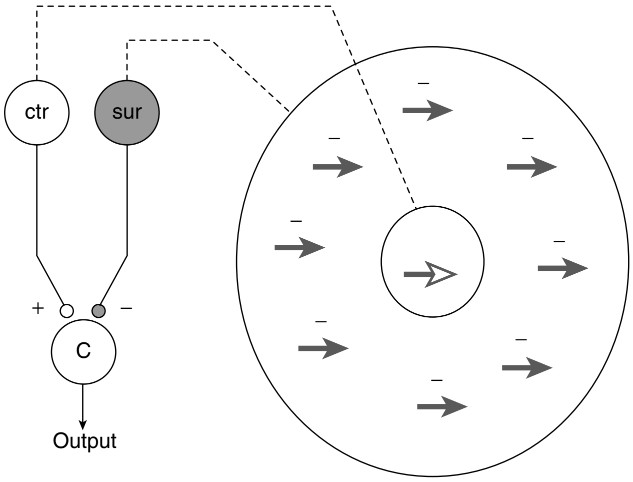


Figure 3 The Circuit for Center-Surround Antagonism

Notes: The center mechanism ("ctr") pools right-detecting EMDs from a small region of the retina, whereas the surround ("sur") pools similar EMDs but over a wider region of the retina. Because the center excites and the surround inhibits the output neuron ("C"), global motion produces no net output.

created by an object that actually moves—such as a running dog—it is always motion with respect to a background. When retinal motion is created by movement on the part of the observer—such as an eye movement—the entire visual world moves in unison. The key distinction here is that object motion is *local* (i.e., only a relatively small part of the visual field moves), whereas self-induced motion is *global* (i.e., the entire visual scene moves). The visual system can detect this difference by performing a comparison of motion between any small region of the visual scene and the larger region in its neighborhood using a circuit known as *center-surround antagonism* (Figure 3). This is a higher order motion processing circuit that involves the integration of responses from a large number of EMDs tiling a region of visual space, and it is found abundantly in regions of the monkey cortex such as MT and MST. One can think of a neuron defining the center mechanism that is activated by a small pool of EMDs responding to the same direction of motion, such as left to right, and a second neuron producing the surround mechanism that pools EMDs of the same preferred direction but over a much larger region of the visual scene. The key is that the center and surround neurons interact in an antagonistic manner so that, when they are both activated, they cancel each other out and produce no net activity. The result is that the center-surround circuit produces activity if there is local motion in its preferred direction, but not if there is global motion (because this activates both the center and the surround) and thus can distinguish motion of objects (local) from that of the self (global).

The other way that the visual system can distinguish self-motion from object motion is to make use of the signals that other parts of the brain generate to produce movement. Consider image motion produced by an eye movement. In order to move the eye in the first place, the brain must generate a command signal that is ultimately sent to the appropriate eye muscles to produce the desired movement. By routing a copy of this signal (called *efference copy* or, alternatively, *corollary discharge*) to the motion sensing circuitry, it can effectively be subtracted from the directional computation and, in this way, eliminated. This mechanism is clearly important for eye movements as evidenced by the fact that we perceive a stable external world despite the fact that our eyes are constantly moving around. If the eye is made to move by external means—such as by *gently* jiggling the eyeball with a finger—the world will appear to jump around quite violently.

Links to Perception

Because of the biological importance and widespread representation of visual motion in the brain, physiological studies of brain circuits for motion have provided some of the deepest insights into the relationship between neurons and perception. These insights have come from experiments pioneered by William T. Newsome and colleagues, in which the "performance" of direction-selective neurons in a monkey's visual cortex was directly compared with the performance of the monkey on a direction discrimination task. They recorded the electrical signals from direction-selective neurons in MT while the monkey performed a perceptual task in which it had to distinguish between two directions, one of which was in the neuron's preferred direction (for example, to the right) and one in the opposite, or null, direction (to the left). On any given trial (consisting of a single stimulus presentation), the motion of a field of dots could be either rightward or leftward and the monkey signaled which direction he saw by making an eye movement to one of two targets. The further wrinkle in the experiment was that the strength of the motion signal could be systematically varied by having some proportion of the dots moving in a particular direction and the rest flickering randomly.

This type of experiment allowed Newsome's group not only to directly compare neuronal and perceptual sensitivity, but also to show how neuronal variability on different trials could actually account for the animals' perceptual decisions. This last point is tricky, but important. On a small proportion of the trials, all of the dots flickered randomly so that there was no net motion in any direction and hence no correct answer. On these trials, the monkey simply had to guess. The interesting result is that, on these guess trials, the electrical signals from the neuron were predictive of the monkey's choices. That is, if the neuron being recorded preferred motion to the right, this neuron tended to be slightly more active on trials where the monkey reported "motion to the right" and slightly less active when he reported "motion to the left." This ability to predict the monkey's choices by listening to a single neuron, referred to as *choice probability,* strongly indicates that the monkey's brain is, in fact, using the signals from this neuron to make the perceptual decision.

If this is really the case, then one should be able to influence the monkey's decisions by artificially activating this neuron. This was done by using tiny electrodes to electrically activate small groups of neurons that had the same preferred direction. The result was as predicted—a spectacular culmination of this line of work. By electrically activating rightward-preferring neurons, the monkey could be biased to report rightward motion even when none was present. Such experiments have established the dominant paradigm for probing the neural basis of perception.

Richard T. Born

See also Animal Motion Perception; Corollary Discharge; Modularity; Motion Perception; Neural Representation/Coding; Self-Motion Perception

Further Readings

Allman, J. M., Miezin, F., & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons. *Annual Review of Neuroscience, 8,* 407–430.

Barlow, H. B., & Levick, W. R. (1965). The mechanism of directionally selective units in rabbit's retina. *The Journal of Physiology, 178,* 477–504.

Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT. *Annual Review of Neuroscience, 28,* 157–189.

Egelhaaf, M., Hausen, K., Reichardt, W., & Wehrhahn, C. (1988). Visual course control in flies relies on neuronal computation of object and background motion. *Trends in Neurosciences, 11,* 351–358.

Livingstone, M. S. (1998). Mechanisms of direction selectivity in macaque V1. *Neuron, 20,* 509–526.

Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience, 21,* 227–277.

MOTION PERCEPTION: SOCIAL

Perceivers are highly sensitive to the movements of objects in the environment. These cues are readily perceived, and such perception is served by distinct neural mechanisms. Yet one class of motion perception appears to be special—the perception of other people moving in the environment, or biological motion. Indeed, the perception of biological motion is an inherently social process. Unique patterns of biological motion support the perception of domains that have been and continue to be of keen importance from a social psychological perspective on person perception, and the basic perceptions of biological motion are susceptible to modulation from high-level social processes. This entry will cover the background, perceiving social categories, identities, and internal states, and how social processes contextualize visual perception as they relate to social motion perception.

Background

The scientific study of biological motion perception enjoys a long history, dating back to the 19th century. Yet research in the perception of biological motion did not take firm hold in the perceptual sciences until the early 1970s when it carved a prominent place. This was largely due to the development of an elegant experimental manipulation that permitted researchers to isolate the body's motion by affixing either lights or reflective material to the major body joints of a person who was filmed in the dark engaging in an activity. When played back to observers, only the lights were visible. Fittingly, such stimuli came to be known as *point-light displays*. With his early reports, Gunnar Johansson set the stage for an explosion of research in the perception of biological motion. Much of the earliest work in this area documented observers' ability to differentiate between human and nonhuman motion or to categorize the action being depicted (e.g., running, walking, jumping jacks). Yet research in the perception of biological motion quickly took a turn to interpersonal judgments.

Building on early observations, researchers became interested in observers' ability to discern personal attributes about the people depicted in point-light displays. Three of these domains, in particular, have been and remain to be important aspects of person perception from a social perspective. These include how the perception of social categories, identity recognition, and emotion recognition are supported by perception of the body's motion.

Perceiving Social Categories

Since Gordon Allport's early work, scholars have recognized the importance of social categorization as a means to facilitate person perception. Once social categories are known, they set forth processes that are difficult to recall. Even when observers do not endorse stereotypes associated with a particular social category, it is difficult to escape the effects of group-based expectations and evaluative tendencies. Indeed, decades of research in social psychology suggest that social categorization is sufficient to arouse stereotypes associated with the category, which subsequently affect attitudes about and behaviors directed toward others. Three major categories tend to dominate social perception—sex, race, and age. Of these, sex categorization has received the most empirical attention in the perception of biological motion.

The body's motion is sexually dimorphic, and observers can reliably identify the sex category membership of people depicted in point-light displays. Some have argued that the motions of men and women are distinct because their body shapes differ. Differences in body shape correspond to distinct motion patterns. By this reasoning, body motion is structurally mediated, and observers' ability to categorize others according to sex reflects an ability to recover the body's form from its motion. The reasons underlying accurate sex perception remain the focus of empirical research, but the finding that observers can discern sex category highlight a privileged role of body motion for the perception of the interpersonally relevant social category, sex.

Perceiving Identities

Social psychological theories of person perception frequently cite two distinct processes involved in person perception. The first, social categorization, was just described. The second, individuation, involves finer discriminations that enable perceivers to identify individuals, not merely the social categories in which they fall. The same logic that led researchers to examine the perception of social categories from biological motion also led some to predict that individual identities might be supported by unique motion parameters as well. Indeed, observers of point-light displays are adept at identifying the body motions of themselves and their friends, presumably due to stable individual differences in body motion. Moreover, when stimuli are generated that exaggerate the spatiotemporal patterns of motion, recognition accuracy is enhanced. Other work has corroborated these effects and has established that recognition ability is due both to perceptual experience with the movements of close others and motor experience with their own actions.

Perceiving Internal States

An ability to determine the internal states of others is crucial for successful social interactions. Of all possible internal states that one might perceive, emotions have been argued to be the most important. Though the majority of work in motion perception has focused on the perception of facial displays of emotion, the body also conveys emotion state. Some have argued that the perception of emotion from the human body can even be more important than perceiving it in the face because, unlike the face, the body conveys not only the appropriate emotional reaction to a situation (e.g., fear), but also the appropriate behavioral response to the situation (e.g., run). Thus, research in the perception of emotional body motion continues to increase. As in other domains of social relevance, observers of point-light emotion displays can accurately categorize the emotion being depicted. Some evidence suggests that perceivers are particularly sensitive to anger and fear displays because of their relevance for personal safety.

Social Processes Contextualize Visual Perception

Recent findings have begun to challenge the long-held assumption that visual processes were independent of higher level processes. Indeed, there is an emerging awareness that visual perception generally, and person perception specifically, is modulated by social cognitive processes. These effects are likely to be the result of interconnections between regions of the brain associated with the visual perception of human action and those associated with the processing of emotional contexts and threat. Additionally, behavioral evidence supports a model of visual perception of human motion that is contextualized by social processes.

Social processes have been shown to contextualize the perception of human motion in at least three ways. First, the very perception of motion varies as a function of the context of the depicted motion. Successive frames of *apparent motion* (such as occurs in movies/film) are perceived differently, depending on the context of the action. When a motion, such as a punch, is directed toward a person in two successive frames (e.g., the beginning point and end point), observers report a strong perception of smooth motion. When the identical motion is directed toward a nonhuman object, such as a refrigerator, reports of perceiving smooth motion are substantially reduced.

Second, social categorization can provide a lens through which body motions are perceived and interpreted. Importantly, the cues that affect judgments of social categories such as sex are abundant and redundant. Some research has shown that social categories are likely to be perceived first from static cues, such as the body's shape. Then, perceived body motion is evaluated in terms of its typicality for the perceived sex. That is, the perception of sex category (male/female) determines how body motion is interpreted. Masculine motions are deemed typical for a man, but not a woman; feminine motions are deemed typical for a woman, but not a man. These assessments of typicality impact judgments of attractiveness and sexual orientation.

Finally, known stereotypes can bias how body motions are perceived and labeled. The emotions anger and sadness, for example, are highly sex stereotyped. These stereotypes affect how perceivers label the sex category of point-light displays that vary in emotion. For example, when observers view point-light displays that depict anger, an emotion category that is stereotypically masculine, observers tend to categorize the target to be men. When observers view displays that depict sadness, an emotion category that is stereotypically feminine, observers tend to judge the displays to be women. In this way, gender-based stereotypes modulate the perception of sex category membership.

Kerri Johnson

See also Motion Perception

Further Readings

Adams, R. B., Jr., Ambdy, N., Nakayama, K., & Shimojo, S. (2008). *The science of social vision.* New York: Oxford University Press.

Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology, 58,* 47–73.

Johnson, K. L., & Tassinary, L. G. (2005). Perceiving sex directly and indirectly: Meaning in motion and morphology. *Psychological Science, 16,* 890–897.

Pollick, F. E., Kay, J., Heim, K., & Stringer, R. (2005). Gender recognition from point-light walkers. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 1247–1265.

MULTIMODAL INTERACTIONS: COLOR–CHEMICAL

Prior to smelling (olfaction) or tasting (gustation) something, we usually see the item producing the smell or taste. This is particularly true when the item is food, which when put in our mouth stimulates taste receptors in the oral cavity and olfactory receptors in the nasal cavity (via what is called retronasal olfaction). It is also often true for items we smell in our external environment (e.g., flowers), which stimulate our olfactory receptors by chemicals inhaled through our nostrils (orthonasal olfaction).

One of the most salient visual properties of an object we are also experiencing through one of our chemical senses (i.e., gustation or olfaction) is its color. Color can serve to predict olfactory or gustatory qualities to follow. For example, before drinking an orange-colored beverage, we might expect it to be a bit sweet and a bit sour (gustatory qualities) and also to be orangey (olfactory quality) because orange-colored beverages have had those properties in the past. Such learned associations between color and both gustatory and olfactory properties can result in an influence of color on odors and tastes. This entry will discuss the influence of color on odor and taste.

Influence of Color on Odor

Certain colors seem to go with or correspond to certain odors. For example, people find that more intense odors correspond to darker colors. Some specific correspondences have been found between particular colors and the odors of compounds used in perfumes. For example, people indicate a correspondence between the smell of cinnamic aldehyde (cinnamon) and the color red, whereas they indicate a correspondence between caramel lactone (caramel) and the color brown. Even complex perfumes have corresponding colors. Perfumes perceived as masculine correspond to blue and perfumes perceived as feminine correspond to pink.

These correspondences are not fully understood. Some might result from learned associations between the color and the odor. For example, after repeatedly seeing an orange-colored solution prior to smelling its orange odor, a person might come to associate the color and odor. Some connections may be innate and reflect a type of synesthesia. How these correspondences might affect people's perception of odors is discussed next.

Visual input, particularly color, helps us identify odors. We have difficulty identifying an odor when the color of the odorant is masked or altered. For example, a cherry-flavored drink colored orange is often misidentified as being orange-flavored (cherry rather than orange being the odor component of the drink). Similarly, red birch beer is often misidentified as cherry soda and white wine colored red is misidentified as red wine. In addition, we are able to identify orthonasal odors both more accurately and more quickly when the color of the stimulus is more appropriate to the odor.

Color can also affect pleasantness by affecting identifiability. For example, grape odor is more often identified correctly and found to be more pleasant when colored purple than when colored yellow. This increase in pleasantness with identifiability might be related to the fact that people find familiar odors more pleasant than unfamiliar ones. Enhanced identifiability contributes to familiarity, thereby increasing pleasantness. Brain areas associated with the hedonic evaluation of odors are more strongly activated by appropriately colored than by inappropriately colored odors.

The presence of a color often appears to signal the presence of an odor. People claim to detect an odor when smelling an odorless Q-tip more often if the Q-tip is colored than if it is not dyed.

Odorous solutions smell stronger when colored (appropriately or inappropriately) than when clear. For example, a strawberry odor smells stronger when colored either red or green than if it is clear. Hiding the color (e.g., by blindfolding) eliminates the enhancement of the odor intensity by the color. Darker colors enhance the perceived intensity of odorous solutions slightly more than do lighter colors.

Enhancement of perceived olfactory intensity by color only occurs for stimuli smelled orthonasally. For solutions smelled retronasally, the effect is reversed. Colored solutions, when put into the mouth smell weaker than equally concentrated colorless ones.

Influence of Color on Taste

Probably because color is less strongly associated with particular tastes than with odors, the effects of color on taste are more variable. Although certain colors seem to correspond to certain tastes, these correspondences appear dependent upon what substance the person is thinking about. For example, a brown color might correspond to a sweet taste if the person is thinking about cola, a bitter taste if he is thinking of coffee, or a salty taste if he is thinking about beef broth. This variability, plus the fact that tastes are easier to identify than smells, might explain the paucity of research on the influence of color on identification and pleasantness of gustatory stimuli.

Many studies (but by no means all) have found that a red color enhances the perceived sweetness of such solutions as plain sucrose, and cherry- and strawberry-flavored drinks. However, red does not seem to lower the sweet (sucrose) taste threshold, although it does increase the bitterness threshold of a caffeine solution.

In addition, blue has been found to decrease the sweetness of a sucrose solution, green to decrease the sweetness of pear nectar (not replicated in a subsequent study), and yellow to increase the sweetness of fruit-flavored solutions. Some studies have found that the presence of *any* color enhances the perceived sweetness of sucrose solutions. Clearly, the influence of color on taste is not as well understood as the influence of color on odor.

Debra A. Zellner

See also Color Perception; Olfaction; Taste

Further Readings

Gilbert, A. N., Martin, R., & Kemp, S. E. (1996). Cross-modal correspondence between vision and olfaction: The color of smells. *American Journal of Psychology, 109,* 335–351.

Johnson, J., & Clydesdale, F. M. (1982). Perceived sweetness and redness in colored sucrose solutions. *Journal of Food Science, 47,* 747–752.

Koza, B. J., Cilmi, A., Dolese, M., & Zellner, D.A. (2005). Color enhances orthonasal olfactory intensity and reduces retronasal olfactory intensity. *Chemical Senses, 30,* 643–649.

Osterbauer, R. A., Matthews, P. M., Jenkinson, M., Beckmann, C. F., Hansen, P. C., & Calvert, G. A. (2005).Color of scents: Chromatic stimuli modulate odor responses in the human brain. *Journal of Neurophysiology, 93,* 3434–3441.

Zellner, D. A., Bartoli, A. M., & Eckard, R. (1991). Influence of color on odor identification and liking ratings. *American Journal of Psychology, 104,* 547–561.

MULTIMODAL INTERACTIONS: NEURAL BASIS

Our sensory world is made up of stimuli in the form of various types of environmental energy. To deal with the vast array of information we are continually bombarded with, specialized sensory systems have evolved, each of which is tuned to a different form of energy (e.g., light, chemical, mechanical). Although each of our sensory systems provides us with a unique perspective on the world, it is only through the synthesis of information between the different senses that we can ultimately form a coherent perceptual unity. Various examples serve to illustrate both the ubiquitous nature of such multisensory interactions as well as the compelling way in which they shape our behaviors and perceptions. Thus, the ventriloquist's act highlights the ability of visual cues (i.e., movements of the dummy's mouth and head) to effectively shift the perceived location of a sound source (i.e., the voice of the ventriloquist). For those who haven't seen the vaudeville act, a similar multisensory "illusion" is evident (or actually, not so evident) whenever we go to a movie theater, where the auditory sources are often far away from the screen. Thus, the actors' voices are typically coming from speakers located far from any actor's position on the screen. Despite the significant spatial mismatch between the visual and auditory cues, we have no trouble in ascribing the appropriate voice to a given character on the screen. One of the most common multisensory perceptual effects happens every time we have a meal, and in which gustatory, olfactory, tactile, and visual information from the food and drink is combined into a wonderful multisensory mélange. In a simpler context, multisensory interactions are also evident when examining behavioral responses, such as simple reaction times, in which participants are asked to press a button as soon as they detect a light or a sound. When the light and sound are presented simultaneously, responses are significantly faster than they are for either of the stimuli when presented alone. Such a speeding underscores the adaptive significance of multisensory processes, and points to the strong evolutionary pressures that likely shaped our ability to make use of redundant (and nonredundant) information from multiple senses.

In order for such multisensory interactions to take place, information from the different sensory systems must come together somewhere within the brain. Indeed, the convergence of inputs from multiple senses takes place at various sites within the central nervous system, ranging from the brain stem to the complex neocortical domains that make us uniquely human. Within these various multisensory brain regions, many individual neurons, the building blocks of the nervous system, receive input from two (or more) sensory systems. Rather than simply serving as passive filters to relay these inputs to the next processing stages, these multisensory neurons actively transform their different sensory inputs, often in ways that give rise to outputs that differ markedly from expectations. This entry will discuss the superior colliculus as a model, multisensory integration, the principles of multisensory integration, multisensory integration from neurons to networks, and beyond the traditional multisensory designations.

The Superior Colliculus as a Model

To date, the best-studied brain structure for examining multisensory interactions at the neural level has been the superior colliculus (SC), a midbrain structure that is situated between the thalamus and the inferior colliculus, a major relay in the auditory system. The SC is an excellent model for these studies for several reasons. First, it contains a rich population of multisensory neurons—neurons that respond to or are influenced by more than a single sensory modality. Second, the SC is organized into an elegant topographic (i.e., maplike) representation of sensory space. Finally, the SC has been long known to play a well-established role in behavior—specifically in guiding movements of the eyes and head toward a stimulus of interest. The importance of the SC as a foundational model for examining multisensory neural processing has come from work showing that its population of multisensory neurons closely resembles those found in other brain regions. Such a result has led to the suggestion that multisensory neurons, regardless of where (or in which species) they are found, abide by a common set of principles as to how they combine and integrate their different sensory inputs. Such commonality may represent an efficient means to "bind" together multisensory neural processes across various brain regions, thus linking changes in behavior to their perceptual correlates.

Multisensory neurons in the SC have discrete receptive fields—areas of sensory space to which they are tuned. A stimulus that is positioned within (or that moves within) the receptive field results in a change in the neuron's activity—it fires more (or fewer) action potentials. Given that they receive convergent input from multiple senses, multisensory neurons have two or more receptive fields, one for each of the senses to which they are responsive. For example, a visual-auditory multisensory neuron will have a visual receptive field and an auditory receptive field. Almost universally, these two receptive fields are closely aligned. Thus, if the visual receptive field is located directly in front of the eyes (i.e., in central space) and slightly upward, the auditory receptive field will be similarly positioned. This spatial registry between receptive fields plays a major role in the construction of the SC's topographic organization and also plays an important role in determining the multisensory interactions exhibited by the individual neuron.

Multisensory Integration

As previously alluded to, at the neuronal level, a multisensory pairing frequently results in a response that differs markedly from what might be predicted by looking only at the neuron's unisensory (e.g., visual alone, auditory alone) responses. Let's take as an example a visual-auditory neuron in the SC. Each time a visual stimulus is moved within this neuron's visual receptive field, two to three action potentials are generated. Similarly, each time an auditory stimulus is turned on within the auditory receptive field, one to two action potentials are generated. Now, in trying to predict the outcome of the multisensory pairing, when the visual and auditory stimuli are presented together, several simple scenarios can be envisioned. In the first of these, the neuron might respond in a way that reflects the more effective of the unisensory responses. Hence, three action potentials would be the expected outcome to each multisensory pairing. In the second scenario, the neuron could be considered as a simple adding machine, with the consequence being an average of three to five action potentials. However, in striking contrast to these predictions, a typical outcome for such a pairing is 10 or more action potentials. Collectively, the transformation that occurs upon stimulus combination has been referred to as *multisensory integration.* Although the previous example illustrates the gains in response seen upon multisensory pairings—collectively referred to as *response enhancements*—multisensory neurons can also exhibit substantially reduced responses under multisensory conditions. These reductions are referred to as *response depressions.* An alternate way to think of the multisensory interaction is in terms of the response predicted on the basis of the unisensory responses. Using this convention, multisensory interactions can be divided into superadditive, additive, and subadditive categories. The circumstances under which these different forms of multisensory interaction can be evoked are described in the following section. The likely utility of this increase (or decrease) in neural activity has been demonstrated by studies that have examined the consequences of the same multisensory pairings on both neural and behavioral responses, and which have shown striking parallels between these responses, strongly suggesting that the activity of these multisensory populations plays an important role in behavioral and perceptual processes.

The Principles of Multisensory Integration

The way in which individual multisensory neurons combine their different sensory inputs has been shown to be critically dependent on several stimulus-related factors. Collectively, these factors have been termed the principles of multisensory integration, because they seem to apply to most multisensory neuronal populations.

The first of these is the *spatial principle,* which simply states that the typical outcome of a multisensory pairing in which both stimuli are presented within their respective receptive fields is a response enhancement. Note that the absolute spatial proximity of the stimuli is not important here—it is simply their relationship relative to the receptive fields. Conversely, when one of the stimuli is presented outside of its receptive field, the end product is either no interaction or a response depression. The presence of such depression has been shown to be contingent upon the presence of an inhibitory region outside of the classical excitatory receptive field.

The second principle, the *temporal principle,* relates to the relative timing of the paired stimuli. In its simplest incarnation, this principle says that the largest relative gain (i.e., enhancement) is seen when the stimuli are presented at or about the same time. As the delay between the stimuli increases, the size of the enhancement declines, until, at long temporal separations, these enhancements may transition to response depressions.

The final principle has been labeled *inverse effectiveness,* in which the relative gain is inversely related to the effectiveness of the individual stimuli in eliciting a response. Thus, the largest response enhancements, like the example previously described, occur when weakly effective stimuli are paired. As the effectiveness of the individual stimuli increases, the relative gain seen upon their combination declines.

Reinforcing the utility of these three basic neural principles is the fact that they have also been demonstrated to apply at the level of human behavior and perception. When thought of in this broader context, these multisensory principles make a great deal of intuitive sense. An environmental event (take as an example a ball bouncing off a table) generally results in visual and auditory cues that are in good spatial and temporal correspondence. As the sources of the visual and auditory energies become progressively more separated in space and/or time, the likelihood that they belong to the same event diminishes. Inverse effectiveness likely reflects the adaptive significance of multisensory processing. Thus, when an event is highly salient, information from a single sensory channel is generally sufficient for its detection and discrimination (think of an explosion in which there is a bright flash of light). In contrast, the major benefits of multisensory integration appear to come when the information provided by any single sensory channel is weak or ambiguous. Under these circumstances, the combined use of information from two (or more) sensory systems greatly facilitates the detection and discrimination of events.

Multisensory Integration From Neurons to Networks

Although the seminal studies of multisensory processing have been conducted at the level of the single neuron, the development of physiological methods that enable the activity of neuronal ensembles (both small and large) to be measured have shown that these populations behave in ways that very much reflect their constituent neuronal contributors. Thus, local field potentials (LFPs), which measure activity changes in local populations on the scale of thousands of neurons, show response enhancements and depressions (and super- and subadditive interactions) in response to multisensory pairings, as do event-related potentials (ERPs) and functional magnetic resonance imaging (fMRI), which measure activity over a much broader spatial scale. Once again, these larger scale multisensory interactions appear to closely abide by the spatial, temporal, and inverse effectiveness principles.

Beyond the Traditional Multisensory Designations

One of the emerging themes in contemporary sensory systems research is that the classical designation of brain regions as being exclusively dedicated to a single sensory modality is now under revision. This is largely a result of recent work that has shown changes in neuronal responses in regions of the visual, auditory, and somatosensory cortex when the "appropriate" stimuli (e.g., auditory stimuli for regions of the auditory cortex) are paired with stimuli from another sensory modality. Typically, these regions of the unisensory cortex are not overtly responsive to this second modality. For example, visual stimuli generally will not evoke responses in the auditory cortex. Rather, multisensory influences manifest as an ability of inputs from the second sense to modulate responses to the "driving" modality. Hence, in the example previously introduced, visual stimuli can alter the responses of neurons in the auditory cortex to auditory stimuli. Similar to the multisensory interactions detailed previously, these alterations can manifest as increases or decreases in activity. The presence of these modulatory influences in each of the major sensory cortices strongly suggests an important role for these cross-modal effects "early" in the sensory processing hierarchies, and has illustrated the importance of multisensory interactions in areas outside of the classic "association" cortices to which these interactions have been historically relegated.

Mark T. Wallace

See also Cross-Modal Transfer; Multimodal Interactions: Visual–Auditory; Multimodal Interactions: Visual–Motor; Neural Representation/Coding; Physiological Approach; Synesthesia

Further Readings

Calvert, G. A., Spence, C., & Stein, B. E. (Eds.). (2004). *The handbook of multisensory processes.* Cambridge: MIT Press.

Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science, 221,* 389–391.

Meredith, M. A., & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology, 56,* 640–662.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses.* Cambridge: MIT Press.

Wallace, M. T. (2004). The development of multisensory processes. *Cognitive Processing, 5,* 69–83.

Wallace, M. T., & Stein, B. E. (1996). Sensory organization of the superior colliculus in cat and monkey. *Progress in Brain Research, 112,* 301–311.

MULTIMODAL INTERACTIONS: OLFACTION–TASTE

See Flavor

MULTIMODAL INTERACTIONS: PAIN–TOUCH

While injecting local anesthetic into a patient's gum, the dentist may grab the patient's cheek and shake it with his or her other hand. This greatly reduces the pain of the injection. The sensory signals produced by shaking, which may be considered a rather vigorous form of touch, are somehow blocking the pain the patient would otherwise feel. Patrick Wall and Ronald Melzack explained this suppression by proposing that pain signals must pass through a neural "gate" in the spinal cord if they are to reach the brain, and that touch signals are sometimes able to close this gate. Later work has broadly confirmed this *gate control theory,* while refining our understanding of how gating works.

Pain gating is possible because of the distinctive anatomy of the sensory pathways for touch and pain. The details of these pathways and their functioning for the perception of pain and touch are discussed in this entry. Separate sensory nerve fibers, called mechanoreceptors and nociceptors, respond to tactile stimulation and to noxious stimulation, respectively. These two types of signals are therefore kept separate from one another as they travel, in neighboring axons, from peripheral tissues to the spinal cord. (Sensory fibers from the face go to the brain stem instead of the spinal cord, but the rules are otherwise the same.)

Once in the spinal cord, the two types of axons follow different paths. Mechanoreceptors split, with a long branch reaching all the way up to the brain, while a short branch terminates in the cord. Nociceptors, on the other hand, *all* terminate in the cord. Therefore, only by being synaptically transmitted to second-order neurons can pain signals reach the brain. As they cross these synapses, pain signals are especially vulnerable to being suppressed or otherwise modified.

Remarkably, the short branches of mechanoreceptors are able to influence pain intensity by interfering with the ability of nearby nociceptors to activate second-order neurons. Mechanoreceptors bring about this interference by activating special inhibitory neurons with short axons. This *segmental inhibition* (so named because it occurs when the tactile and noxious stimuli are close enough together for their signals to interact within individual or adjacent cord segments) is the primary physiological process underlying pain gating.

We make use of pain gating in everyday life, when we gently rub a bruise, press our forehead during a headache, or (because itch is closely related to pain) scratch an insect bite. But research has shown that vibration, as when the dentist shakes a patient's cheek, is the most effective form of touch for closing the gate. Vibration activates mechanoreceptors vigorously because they are stimulated anew by each cycle; their messages reach the central nervous system and strongly prod inhibitory neurons to block pain messages. Drugstore vibrators relieve an aching back or shoulder through a combination of pain gating and their effect on muscle tension.

Pain can also be suppressed in other ways. For example, a mild pain may be reduced or eliminated when the person has a more intense pain in another part of the body. This phenomenon is called *diffuse noxious inhibitory controls,* or DNIC. An example is that a person with arthritis in several joints may note that his or her pain seems to "travel around the body" from day to day, affecting primarily one joint at a time. Emotional state can also affect pain. In one recent case, a woman severely cut her hand. As she drove herself to the emergency room, bleeding profusely and fearing for the consequences, she felt no pain. Only when the doctor assured her that she would be fine did severe pain envelop her hand.

In both of these situations, structures in the brain stem send inhibitory signals down the spinal cord to the gate, where much of the actual suppression occurs. In the case of DNIC, the brain stem structures are activated by intense pain messages that have already passed through the gate; in the case of emotional influences, the brain stem structures are activated by emotion-related messages from other parts of the brain.

Although pain gating has been the object of scientific study for half a century, the converse phenomenon, *touch gating,* has only recently been discovered. Touch gating is the suppression or compromising of touch by pain. It can be quantified by determining the detection threshold for vibration (i.e., the minimum amplitude at which a vibration can be felt), at a particular place on the skin, under baseline conditions and then again when a painful stimulus is applied nearby. In a classic study by Vania Apkarian and his colleagues, the noxious stimulus was hot water, flowing through tubes in a metal ring pressed against the skin surrounding the spot touched by the vibrator. As the ring's temperature rose and became painful, vibration threshold increased as well. In some chronic pain conditions, vibration threshold is elevated in the painful region even when no noxious stimulus is applied by the experimenter: Clinical pain, like experimental pain, thus appears capable of producing touch gating.

Some psychophysical research indicates that in touch gating, the gate is located in the brain, not the spinal cord. This is consistent with anatomy, because touch signals, as described earlier, can travel all the way from the skin to the brain without having to cross a synapse. Once they are in the brain, they become susceptible to interference. This localization of touch gating in the brain, perhaps even the cortex, has led some to suggest that it may reflect "merely" an attentional shift: distraction of the subject from the weak vibration by the much more salient pain. Evidence on this question is inconclusive, but it is worth noting that there is nothing insubstantial about attention, which like all psychological phenomena, depends on physiological processes.

Mark Hollins

See also Pain: Cognitive and Contextual Influences; Pain: Physiological Mechanisms; Perceptual Development: Touch and Pain

Further Readings

Apkarian, A. V., Stea, R. A., & Bolanowski, S. J. (1994). Heat-induced pain diminishes vibrotactile perception: A touch gate. *Somatosensory and Motor Research, 11,* 259–267.

Bini, G., Cruccu, G., Hagbarth, K.-E., Schady, W., & Torebjörk, E. (1984). Analgesic effect of vibration and cooling on pain induced by intraneural electrical stimulation. *Pain, 18,* 239–248.

Hollins, M., & Sigurdsson, A. (1998). Vibrotactile amplitude and frequency discrimination in temporomandibular disorders. *Pain, 75,* 59–67.

Melzack, R., & Wall, P. D. (1965). Pain mechanisms: A new theory. *Science, 150,* 971–979.

MULTIMODAL INTERACTIONS: SPATIAL PERCEPTION IN TOUCH AND VISION

Spatial perception is essential for locating objects, distances and directions. It is more closely associated with vision than the other senses in humans, as in many primates. Evidence for spatial perception by active touch has been growing from the early twentieth century onward.

Tests with touch highlight the reference relations that characterize spatial perception. We cannot feel the length of a piece of string, except in relation to its start and endpoint. The combination of touch with scanning and exploring movements—now usually called haptic perception—makes it more obvious that to feel—or see—where an object is (locating it), how far (distant) it is, and in what direction it lies depends crucially on perceiving it in relation to some reference anchor or boundary. This entry discusses the integration of inputs from internal and external sources, reference organizations in spatial vision and in active touch without vision, combining and deleting reference cues and boundary points in shape illusions, and neural and brain activities in spatial perception.

Integration of Inputs From Internal and External Sources

Spatial perception depends on the integration of stimuli from the external environment and stimuli arising within the body of the perceiver. The resulting organizations afford the reference cues that are needed to specify locations, distances, or directions for perception and action in different task and stimulus conditions.

Two main types of reference organizations have been distinguished: Seeing or feeling where your purse is relative to other objects, or in relation to the surrounding edges of the table on which it lies, shows that you located it by reference to external cues. Seeing or feeling where the purse is relative to your hand position or body posture shows reference relations based on cues from within the body. External and body-centered reference relations are considered briefly in spatial vision and in active touch without vision.

Reference Organizations in Spatial Vision and in Active Touch Without Vision

Spatial vision is typically associated with external reference cues. We see objects relative to each other and to the boundaries of visual scenes, and can find them by reference to these cues. It is less well-known that spatial aspects of vision also depend critically on (proprioceptive) cues from head and neck postures, and interact with gravitational and directional stimuli from specialized sense organs in the (inner ear) vestibular system. The advantage of spatial vision is that body-centered cues for the upright—vertical—direction are usually congruent with vertical "top-bottom" features of the other gravitationally oriented objects in the external environment.

Relations between external cues are rarely experienced in the total absence of visual experience. By contrast, touch and movement (kinesthetic) cues in locating objects by active touch relate consistently to cues from body and limb postures. Blind children who depend on haptic perception for many tasks in tabletop spaces tend to rely on such body-centered reference. That has been found with tests that disrupt these relations by rotating the display, or by changing the perceiver's body posture relative to it. Relying on body-centered reference relations makes it more difficult to infer new directions.

Nevertheless, total lack of vision from birth does not prevent people from becoming excellent chess players, or from making complex spatial inferences. It suggests that they have access to integrated spatial information. Effects of adding and deleting external cues in haptic perception and of a common discrepancy in shapes that produce the same illusion in vision and touch are considered next.

Combining and Deleting Reference Cues and Boundary Points in Shape Illusions

Explicit instructions on how to use an actual external surrounding frame in conjunction with body-centered cues greatly improved haptic perception. Blindfolded participants used one hand to explore the locations of "landmark" shapes along a raised (embossed) route in a tactile map. Instructions to scan the surrounding square frame with the other hand, and to relate the position of the landmarks to the frame, halved the number of errors in locating the landmarks, although the display had been rotated to disrupt body-centered reference. Accuracy doubled when participants were told to use intact body-centered cues concomitantly with external frame cues. Inexperienced participants evidently needed explicit instructions to coordinate frame and body-centered cues in active touch, despite years of normal vision. Initial explicit instructions on combining the two types of reference cues could also help people without visual experience because most activities become "automatic" with familiarity and practice. But that needs additional tests.

Another study tested the effects of added external cues on haptic perception in spatial tasks further by using spectacles that simulate different types of impaired vision. As expected, adding full vision greatly improved the accuracy of haptic perception of locations in a tactile map. Adding two different types of residual vision reduced, but did not eliminate, the advantage over touch alone. By contrast, adding vision that consisted only of diffuse light perception, which excludes all shape and spatial cues in tabletop space, did not differ in accuracy at any location from touch without added vision. The findings reinforce evidence that the advantage of spatial vision over touch depends on added reference cues.

T-shapes produce the same length illusion in vision and touch. The shapes consist of two lines of equal length in which one line bisects the other at right angles. The continuous bisecting line is seen and felt as longer than the bisected line. Further tests suggested that the junction point on the bisected line is perceived as a boundary that bisects the line into two smaller lengths, so that continuous line appears longer.

Müller-Lyer shapes consist of a shaft that is bounded by small, angled wings. (See Cultural Effects on Visual Perception, Figure 1.) Shafts that are bounded by diverging wings are seen and felt as much longer than the same-size shafts with inward converging wings. The angled wings produce discrepant length cues relative to the shaft. Explicit instructions to use body-centered reference cues almost eliminated the illusion in both touch and vision, suggesting that boundary points that produce discrepant length cues are a common factor in the two length illusions in vision and touch.

Neural and Brain Activities in Spatial Perception

Inputs from vision, touch, and movement feed into distributed areas of the brain that are activated in different spatial tasks via complex connecting neural pathways at different levels in the system. The parietal area functions prominently in the integrative activities that yield coherent spatial perceptions. It lies between frontal (cognitive) and the visual regions at the back of the brain. Inputs from tactile and motor areas come together in a deep fold between frontal and parietal areas. Different cell populations are activated by external and body-centered reference cues, but a few cells in adjacent cell populations are activated in both conditions.

The complex interactions of neural paths that form connecting links and circuits have been described as interrelated neural networks. The description applies to the behavioral data researchers have so far on how spatial processing takes place with peripherally diverse inputs from vision and touch.

Susanna Millar

See also Attention: Spatial; Haptics; Multimodal Interactions: Visual–Haptic; Vision

Further Readings

Levine, D. S. (2002). Neural network modeling. In H. Pashler & J. Wixted (Eds.), *Stevens' handbook of experimental psychology* (Vol. 4, pp. 223–269). New York: Wiley.

Millar, S. (2008). *Space and sense.* New York: Psychology Press.

Paillard, J. (Ed.). (1991). *Brain and space.* Oxford, UK: Oxford University Press.

Snyder, L. H., Grieve, K. L., Brotchie, P., & Anderson, R. A. (1998). Separate body and world–referenced representations of visual space in the parietal cortex. *Nature, 394,* 887–891.

MULTIMODAL INTERACTIONS: TACTILE–AUDITORY

When we explore a texture with our hands or just rub our hands, the tactile percept is accompanied by simultaneous sounds. Although not always obvious in everyday life, this auditory feedback conveys useful information about surface properties. This entry describes examples of how auditory feedback influences tactile perception, and some of the proposed mechanisms. Experiments can unmask such an interaction between auditory and tactile perception, for example in what Jousmäki and Hari called the *parchment-skin illusion,* experimentally demonstrated as follows: subjects rubbed their palms together while the ensuing sound was played back to them in headphones. When the high frequencies of the sound (above 2,000 hertz, Hz) were selectively enhanced, subjects reported that the palmar skin feels dry and parchmentlike, compared with trials when the original sound was played back. This striking example of audiotactile interaction highlights an aspect of multisensory perception that has been studied much less frequently than audiovisual interaction.

In another example, sounds combined with vibrotactile stimuli applied to the fingers were reported as louder than the same sounds presented alone. This effect was demonstrated using vibrotactile stimuli applied via a tube that subjects grasped, resulting in stimulation of their fingers and palms with a perceived intensity of 24 to 28 decibels (dB) above threshold. Tones were 10 dB above threshold, measured within masking noise. Both stimuli were of the same frequency, 200 Hz. Subjects adjusted tones to match a reference tone and chose lower intensities in trials with simultaneous auditory and tactile stimulation, compared with auditory-only stimulation. To demonstrate audiotactile interaction, this experiment kept stimuli at low intensity levels, in line with the *inverse effectiveness rule:* multisensory interactions are most evident when at least in one sensory modality the stimulus is of low intensity, making information from that modality alone unreliable.

It is widely accepted that two further rules govern the size of multisensory interaction effects: Interaction is most likely to be evident when stimuli occur simultaneously (*temporal rule*) and when they are perceived as arising from the same spatial location (*spatial rule*)—conditions that would typically be met when stimuli represent a single audiotactile object. Whether or not the spatial rule applies to audiotactile interaction has been questioned in recent studies. Moreover, it is controversial whether the three rules of multisensory integration, mostly based on experiments at the single-neuron level, fully apply to behavioral studies.

Which of the two senses is dominant when information from both modalities, auditory and tactile, is available to the subject? The weighting depends on several factors, one of them being the subject's long-term experience of how useful information from a given modality is for a certain task. However, subjects can modify the weighting depending on task demands: It has been reported, for example, that subjects who use a probe to explore a texture weight tactile information stronger than during exploration with bare fingers. This adjustment reflects the louder sounds and the lower reliability of tactile feedback during exploration with a probe.

What may be the neural basis of audiotactile interaction that—as the previous examples illustrate—works both ways, with auditory stimuli biasing tactile perception and vice versa? One might consider two interesting parallels between the two modalities, as pointed out by von Békésy nearly 50 years ago: First, they both require neural processing with analyses of rich temporal structures, for example, when sound or vibration frequencies are discriminated. Second, the brain processes auditory as well as vibrotactile stimuli as input from receptors that transduce movement into neural activation: from auditory hair cells that are excited when sound waves make the basilar membrane of the inner ear flex, and from the rapidly adapting Pacinian corpuscles in the skin that are most sensitive to vibration in the range of 100 to 300 Hz—frequencies that are well within the sensitivity range of the human auditory system. So could it be that the auditory cortex, highly specialized in the analysis of temporal patterns (as they occur in speech and environmental sounds), contributes to the processing of temporally structured vibrotactile stimuli?

This question has been addressed in a number of studies using noninvasive neuroimaging, namely functional magnetic resonance imaging and magnetoencephalography. Complementary in methods, these studies found that tactile stimuli can activate the auditory cortex (more precisely the "belt" areas that surround the primary auditory cortex). This cross-modal activation might reflect that the analysis of the temporal structure is supported by the auditory cortex. Indeed, the auditory cortex activation has been reported when a deaf person received vibrotactile stimuli to the hands and discriminated standard versus deviant frequencies.

Audiotactile interaction is also relevant to research on sensory substitution, where stimuli of one modality can compensate for sensory deficits in another modality, for example to support speech perception by hearing-impaired individuals. Scientific literature from the 1980s and 1990s specifies devices that convert auditory stimulus properties (such as frequency and/or intensity and the respective temporal pattern) into tactile stimuli applied via an array of tactors, for example to the abdomen. However, it is unclear whether such devices are widely used. In any case, audiotactile interaction is promising for future applications where noise that affects just one modality needs to be compensated for.

Martin Schürmann

See also Audition; Cutaneous Senses; Surface and Material Properties Perception; Texture Perception: Tactile

Further Readings

Caetano, G., & Jousmäki, V. (2006). Evidence of vibrotactile input to human auditory cortex. *Neuroimage, 29,* 15–28.

Jousmäki, V., & Hari, R. (1998). Parchment-skin illusion: Sound-biased touch. *Current Biology, 8,* R190.

Lederman, S. J., & Klatzky, R. L. (2004). Multisensory texture perception. In G. A. Calvert, C. Spence, & B. R. Stein (Eds.), *The handbook of multisensory processes* (pp. 107–122). Cambridge: MIT Press.

Schürmann, M., Caetano, G., Hlushchuk, Y., Jousmäki, V., & Hari, R. (2006). Touch activates human auditory cortex. *Neuroimage, 30,* 1325–1331.

MULTIMODAL INTERACTIONS: THERMAL–CHEMICAL

Recall an experience in which you consumed a meal that was fresh from the oven and served too

Table 1 Thermosensitive Receptor Proteins and Their Thermal Activation Range

Protein name	*Proposed thermal sensitivity range (°C)*
TRPA1	4–17
TRPM8	8–28
TRPV4	25–34
TRPV3	33–37
TRPV1	43 and above
TRPV2	52 and above

hot. Your mouth may have felt like it was burning; you may have even begun to sweat. Now recall an experience in which, on a dare, you ate a raw chili pepper. In that situation, you may also have experienced a burning mouth and a sweaty forehead. Interestingly, these same responses are evoked by physically different stimuli. In the case of the meal being served too hot, the stimulus is one of a thermal nature—heat. In the case of the chili pepper, the stimulus is a chemical. How can these two, disparate stimuli evoke the same burning perception and the same physiological response (sweating)? The answer to this question is beginning to emerge due to recent advances in molecular biology and neuroscience. This entry will cover thermal and chemical sensitivity, and the perception of chemical-thermal interactions.

Thermal Sensitivity

The ability to sense and respond to environmental temperatures is critical for human survival. As external temperatures increase or decrease, thermal changes are detected by receptors in the cell membranes of *thermosensitive* nerve cells. These cells are found throughout the body, including peripheral tissues like skin and muscle, as well as in the mucous membranes of the oral and nasal cavities. Detection of temperature changes by these receptors causes a rapid excitation of the nerve cell, and this signal is conveyed to the brain, where it is ultimately interpreted as heat or cold. However, as temperatures (measured here in Celsius, C) exceed a certain threshold and become too hot (~41 °C) or too cold (~4 °C), the stimulus is additionally perceived as painful. It is now known that thermal sensitivity is due to the existence of specific protein receptors that are tuned to respond to specific temperatures. To date, six receptor proteins that respond to thermal stimuli have been identified. Each of these proteins responds to a specific temperature range. For instance, the protein TRPM8 (transient receptor potential melastatin-type 8) has been shown to respond to cool temperatures between 8 to 28 °C, whereas a different protein, TRPV3 (transient receptor potential vanilloid-type 3), has been shown to respond to warm temperatures between 32 to 37 °C. Collectively, these proteins respond to temperatures that vary in range from 4 to 53 °C and form a sort of physiological thermometer. Table 1 lists all of the known protein receptors that respond to thermal stimuli and the temperature ranges that activate them.

The receptor proteins that respond to temperatures are found in two types of temperature-sensitive cells—those that respond to nonpainful temperatures and those that respond to painfully hot or cold temperatures. Cells that respond to painful stimuli are called *nociceptors* and express the receptor proteins that are activated by potentially harmful temperatures (TRPA1, TRPV1, TRPV2). Cells that respond to nonpainful thermal stimuli can be subdivided into cold-sensitive and warm-sensitive populations. Cold-sensitive cells exclusively express receptor proteins that are activated by cool temperatures (TRPM8), whereas warm-sensitive cells express proteins sensitive to warm, but not hot, temperatures (TRPV3, TRPV4). Although this organization provides the molecular and neural basis for temperature perception and discrimination, it fails to explain how chemicals can similarly evoke heat, cold, and pain sensations.

Chemical Sensitivity

Interestingly, the same receptor proteins that respond to thermal stimuli also have the capacity to respond to some chemical stimuli. However, as chemical and thermal stimuli are physically distinct, it is likely that the mechanisms by which these receptors are activated are also distinct. To date, a variety of natural and synthetic molecules have been identified that can activate thermosensitive receptors. The most well-studied protein

Table 2 Thermosensitive Receptor Proteins and Their Chemical Ligands

Protein name	*Natural chemical ligands*
TRPA1	thiocyanates, cinnamaldehyde
TRPM8	menthol, eucalyptol
TRPV4	bisandrographolide, arachidonic acid
TRPV3	camphor, thymol, eugenol
TRPV1	capsaicin, piperine, gingerols, marine toxins
TRPV2	Δ9-THC

receptors in this regard are TRPV1, which, in addition to painful heat, also responds to capsaicin (the pungent principal in chili peppers) and TRPM8, which responds to both cool temperatures and menthol. Table 2 provides a list of the known thermosensitive receptors and natural chemicals that activate them.

When a chemical stimulus binds to its cognate receptor protein, the nerve cell expressing that receptor becomes excited. As with thermal stimuli, this activation is conveyed to the brain, where it is ultimately interpreted as painful or nonpainful heat or cold.

Human Perception: Chemical–Thermal Interactions

Evolutionary pressures have selected for receptors that are sensitive to a broad range of temperatures and as such, the "normal" stimuli for these receptors are thermal. Similarly, the thermosensitive neural pathways have evolved to deliver to "decision-making" centers of the brain key environmental information that increases likelihood for survival. Therefore, when activation of these pathways is by abnormal sources (e.g., chemicals), the activity conveyed to the brain is still interpreted as thermal information. Thus, menthol activation of TRPM8 activates cold-sensitive pathways that convey signals to the brain interpreted as cold, whereas capsaicin activation of TRPV1 initiates signaling in pain pathways that is ultimately perceived as burning pain. When thermal and chemical stimuli are simultaneously present, interactions can occur that modulate the sensation that is ultimately perceived. For instance, most of us have experienced the extra-strong cooling sensation that is evoked when sucking on a strong mint and concurrently drawing air into the mouth. Similarly, a nice big drink of ice cold water puts the fire out when eating spicy food. These everyday examples of multimodal thermal–chemical interactions have been substantiated in controlled scientific studies.

Nature has conceived of a way for humans to sense and respond to environmental temperatures. The specific thermal sensitivity of thermosensitive receptor proteins provides the molecular basis that allows humans to detect and differentiate thermal stimuli. The fact that these same receptors can also be activated by specific chemicals explains why some natural plant products evoke a seemingly paradoxical thermal sensation. However, we should be thankful for this multimodal sensitivity because it is hard to imagine a world without the refreshing sensation of chewing gum or the biting pungency of piquant salsa.

Christopher Simons and Jay Slack

See also Common Chemical Sense (Chemesthesis); Pain: Physiological Mechanisms; Temperature Perception

Further Readings

Appendino, G., Minassi, A., Pagani, A., & Ech-Chahad, A. (2008). The role of natural products in the ligand deorphanization of TRP channels. *Current Pharmaceutical Design, 14*(1), 2–17.

Dhaka, A., Viswanath, V., & Patapoutian, A. (2006). TRP ion channels and temperature sensation. *Annual Review of Neuroscience, 29,* 135–161.

MULTIMODAL INTERACTIONS: VISUAL–AUDITORY

Humans are often exposed to visual and auditory information that arises from objects and events in their environment. For example, imagine that you are watching a game of tennis on television and someone accidentally mutes the sound. You will probably notice that following the game becomes harder, not just because the narration is no longer available, but also because perceiving the timing of the impact of the ball on the ground and racket is harder. Now, imagine that instead of watching a game, you are playing tennis yourself. Plugging the ears would strongly interfere with your ability to play because now not only can you not perceive the timing, speed, and location of the ball as accurately, but you also cannot coordinate your actions accordingly. The human nervous system has evolved and acquires ways of utilizing the correlated visual and auditory information for achieving a more accurate and reliable perception and action in the environment. This entry covers the auditory–visual interactions in perception, auditory–visual illusions, and auditory-visual interactions in memory and learning.

Auditory–Visual Interactions in Perception

Although we are almost never consciously aware of it, the interplay between auditory and visual modalities is always operating in daily life. The reason we are not consciously aware of the interactions between vision and hearing is that under normal circumstances, the sights and sounds that correspond to the same object convey consistent information; for example, the sound of the tennis ball hitting the ground and the image of it both agree in the time and location of impact (as well as other attributes, such as speed, weight, etc.). Therefore, our overall estimate of the time and location of impact is a unified and coherent one.

One may ask, how could the two estimates be anything but consistent if they arise from the same object? The answer is that estimates in each sensory modality are always corrupted by noise, noise in the environment (e.g., fog affecting the rays of light, or clutter affecting the sound waves) and noise in the brain (the firing of the neurons is noisy). Therefore, even the same exact stimulus can elicit different neural responses at different times. Therefore, the same tennis ball hitting the same location x on the ground time after time may be heard at location x in one time and at location x + 5 centimeters at another time, and likewise for visual perception. Therefore, even stimuli that arise from the same event can be slightly inconsistent in the sensory estimates they invoke in the nervous system.

The reason we are not aware of such inconsistencies is that the nervous system fuses the signals into one unified estimate by combining the estimates according to their respective reliability. For example, if the tennis court is well lit and the observer has normal visual acuity, then visual estimate of location is likely more reliable than the auditory estimate, as the auditory spatial resolution is generally not as good as vision, and therefore the overall estimate of location will be largely biased toward the visual estimate, while also influenced by the auditory estimate. On the other hand, if it is nighttime, and the court is not well lit or the observer has bad eyesight, then the visual estimate may be less reliable than the auditory estimate, and then overall estimate of location may be determined primarily by the auditory information. The same principle applies to the timing of the bounce. Under normal circumstances the auditory estimates of time are generally more reliable than visual estimates, and therefore dominate the overall estimate of time, but if there is a lot of noise in the background and the auditory estimate is not reliable, then the perception of time may be primarily determined by vision.

Auditory–Visual Illusions

Much of the knowledge of cross-modal interactions has been obtained through experiments that induce a conflict between two modalities and probe the observer's perception. For example, it has been found that if sound is presented at a location that is moderately different from the location of a visual stimulus, it is often perceived to be originating from the same location as the visual stimulus. This effect is known as the *ventriloquism effect,* and is the same effect that ventriloquists have exploited for centuries for their puppet shows. It is also the same effect that we all experience every time we watch TV or a movie at the movie theater, where the voice of the actors is perceived to originate from the same location as the image of the actors on the screen, as opposed to the fixed location of the speakers. This occurs because the visual estimates of location are typically more accurate than the auditory estimates of location, and therefore the overall percept of location is largely determined by vision. Conversely, perception of time, wherein auditory estimates are typically more accurate, is dominated by hearing. One example of this is perception of a number of pulsations (which largely involves temporal processing). If a single flash of light is accompanied by two or more beeps, observers often perceive multiple flashes as opposed to a single flash. In this case, sound dominates the perception as it is generally more accurate in this task. This effect is known as *sound-induced flash illusion.* Another intriguing demonstration of auditory–visual interactions is in the domain of speech perception, and it is known as the *McGurk effect.* If the video of an individual articulating the syllable /ga/ is played synchronously with the sound of an individual saying the syllable /ba/, the syllable /ba/ is often perceived as /da/. This reveals the strong auditory–visual interactions that take place during speech perception, which appears as a purely auditory task.

Auditory–Visual Interactions in Memory and Learning

If auditory-visual interactions are ubiquitous in perception, can they also play a role in learning and memory? Indeed, recent studies have shown that auditory–visual interactions can facilitate memory and learning. For example, it was found that observers were able to recognize the image of objects that were previously presented accompanied by their corresponding sound better than the image of objects that were initially shown only visually. Even more surprisingly, learning of a visual task was recently shown to be facilitated by congruent sounds during training. Observers learned to detect visual motion much faster and much better when, during training, the visual motion was accompanied by auditory motion, despite the fact that sound was absent during testing. Therefore, sound appears to help with the encoding and/or retrieval of the visual information. Conversely, visual information has been shown to facilitate auditory learning. Training with voices that are paired with video clips of talking faces is more effective in inducing learning of voices than training with voices alone, even when videos are absent during testing.

Ladan Shams

See also Bayesian Approach; Binding Problem; Multimodal Interactions: Neural Basis; Neural Representation/Coding; Perceptual Learning

Further Readings

Ernst, M. O., Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8,* 162–169.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences, 27,* 712–719.

Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: Plasticity and interactions. *Current Opinion in Neurobiology, 11,* 505–509.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses.* Cambridge: MIT Press.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88,* 638–667.

MULTIMODAL INTERACTIONS: VISUAL–HAPTIC

Until recently, most textbooks on human perception considered each of the senses (e.g., vision, hearing, touch, olfaction, and taste) in isolation, as if each represented an independent perceptual system. However, in most situations our senses receive correlated information about the same external objects and events, and this information is typically combined by the brain to yield the rich multisensory percepts that fill our everyday lives. This entry highlights one important aspect of multisensory perception, namely, how what we see can influence our perception of haptically explored objects and surfaces. Haptic perception (i.e., tactile perception that involves active, as opposed to passive, touch) provides us with information concerning both the substance (hardness, weight, temperature, texture, etc.) and structural properties (size, shape, and volume) of the objects with which we interact.

Dominance

Traditionally, philosophers believed that touch dominated over, and even educated, vision. However, over the last 75 years or so, psychologists have conducted many studies showing that vision frequently dominates over touch. For example, J. J. Gibson reported that when people ran their fingers up and down a straight rod, they perceived it as being curved if they simultaneously looked through lenses that made the rod look curved. As soon as the participants closed their eyes, however, the rod felt straight again. Similar results were reported by Irvin Rock and his colleagues in a now-classic series of experiments in which participants rated their impression of the size of a small object that they could either see (through a distorting lens) or feel, or both see and feel at the same time. This kind of visual dominance over the perceived size and shape of the haptically explored objects is so strong (and automatic) that it cannot easily be overridden by instruction. In fact, Irvin Rock and Charles Harris went so far as to state that vision completely dominates touch and even shapes it.

Taken together, these and many other published results subsequently showed that vision typically dominates over the haptic perception of both the substance and structural properties of objects when the senses are put into some kind of intersensory conflict. Researchers have also demonstrated that visual dominance effects tend to be more pronounced when the stimuli presented in the two modalities originate (or at least are perceived to originate) from the same spatial location at more or less the same time. For example, Sergei Gepshtein and colleagues showed in 2005 that visual and haptic cues were combined more effectively when they originated from the same spatial location (rather than from different locations), whereas other researchers have shown that people can still integrate visual and haptic information in a near-optimal manner when they come from different locations, just as long as participants believe that what they are seeing and feeling refers to the same object (as when one looks in a mirror in one location in order to see an object that is being haptically explored at a different location). Epp Miller also reported in 1972 that an observer's beliefs concerning whether visual and haptic sensory impressions belonged together, a phenomenon known as the *unity assumption,* also modulate the extent to which the visual cues influence haptic form perception.

Theories

Over the years, various theories have been put forward to try to account for the apparent ubiquity of the visual dominance effects observed in the laboratory, including the modality-appropriateness hypothesis and the directed attention hypothesis. According to the directed attention hypothesis, people tend to direct their attention more toward visual inputs in order to compensate for the poor alerting, or arousing, qualities of visual stimuli. This attentional bias results in (attended) visual inputs (i.e., those sensory impressions that people are concentrating on) being weighted more heavily than those from the other relatively "less attended" sensory modalities (e.g., touch). By contrast, according to the modality-appropriateness hypothesis, our brains tend to favor information from the sense that is most appropriate for the task at hand. The argument is that visual information frequently dominates because vision is the sense that normally provides the most accurate information concerning the judgment being made (at least for the kinds of perceptual judgments that psychologists are fond of asking their participants to make). Consistent with this latter view, researchers investigating the multisensory perception of surface texture have shown that both visual and tactile (or haptic) cues contribute to people's perception of the felt texture (or roughness) of an object's surface, with the extent to which one sense is preferred over another depending on the nature of the task being performed, and the particular surface being evaluated.

Marc Ernst and Marty Banks brought some much-needed mathematical rigor to the field of sensory dominance research recently by showing that maximum likelihood estimation can provide an excellent quantitative account of the integration of visual and tactile/haptic cues in a task in which participants had to judge the height of a bar that they could see and also feel between the thumb and index finger of one hand. Ernst and Banks showed that adding noise to the visual signal (presented via computer) resulted in their participants increasingly relying on haptic information when making their judgments (i.e., they were able to show that visual dominance was not hard wired). According to the maximum-likelihood estimation account of sensory dominance, the human brain combines sensory inputs in a manner that is very close to that of a statistically optimal multisensory integrator. That is, the multisensory integration of disparate unisensory inputs appears to maximally reduce the uncertainty of (or variance associated with) our multisensory estimates of external stimulus qualities (given that all sensory estimates are intrinsically noisy). Subsequent research has shown that the maximum-likelihood account provides a surprisingly good account of the relative contribution of each of the senses to multisensory perception in a variety of different settings and for a variety of different combinations of stimulus modalities. Nevertheless, there may still be some residual role for directed attention in explaining sensory dominance.

Further Research

Our awareness of the objects in the world around us is determined by a constant interplay between vision, haptics, and the other senses (e.g., audition, olfaction), and as such, multisensory integration is now (rightly) considered the norm, rather than the exception, in perception research. A large body of cognitive neuroscience research currently supports the view that our tactile/haptic perception of both the structural and surface properties of objects is profoundly influenced by what we see, when touching, interacting with, and/or evaluating them. One outstanding question in this area that has yet to receive a satisfactory answer is whether there are any substantive individual differences (perhaps attributable to differences in practice or expertise) in the way in which people integrate haptic and visual information. Are there, for example, individual differences in the extent to which vision dominates over tactile/haptic perception? Only further research will tell.

Charles Spence

See also Action and Vision; Cutaneous Perception; Haptics; Reaching and Grasping; Vision

Further Readings

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429–433.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8,* 162–169.

Rock, I., & Harris, C. S. (1967, May 17). Vision and touch. *Scientific American, 216,* 96–104.

Spence, C., & Gallace, A. (2008). Making sense of touch. In H. Chatterjee (Ed.), *Touch in museums: Policy and practice in object handling* (pp. 21–40). Oxford, UK: Berg.

MULTIMODAL INTERACTIONS: VISUAL–MOTOR

See Action and Vision; Perceptual-Motor Integration

MUSIC COGNITION AND PERCEPTION

Music cognition and perception is the scientific study of those mental and neural operations underlying music listening, music making, dancing (moving to music), and composing. It is intrinsically interdisciplinary, drawing on methods from cognitive and sensory psychology, neuroscience, musicology, computer science, music theory, and sociocultural aspects of music, with genetics and evolutionary biology becoming increasingly relevant. Music processing is a complex, higher cognitive activity engaging many areas of the brain and employing many distinct cognitive operations. As such, music has revealed itself to be a useful window into understanding functions of the mind and brain and informing large issues in cognitive psychology such as memory, attention, perceptual organization, categorization, and emotion. This entry describes the building blocks of music; musical structure, grammar, syntax, and semantics; memory for music and musical imagery; ability, disability, genetics, "talent," and musicianship; emotion and expectation; hemispheric specialization; and music preferences and individual differences.

Building Blocks of Music

The basic elements of any sound are loudness, pitch, duration, timbre, spatial location, and reverberation. These dimensions are separable in that each can be varied without altering the others, allowing the scientific study of one at a time. Of the six, *pitch* and *loudness* are psychological constructions that map loosely (and perhaps nonlinearly) to the physical dimensions of *frequency* and *amplitude*.

When more than one tone is present, the sequence of pitches defines a musical *interval,* and intervals define *contour*—the direction of movement in a sequence of tones (up, down, or the same) without regard to the size of the intervals. Contour may be subject to preferential processing—infants attend to it more readily than they do intervals, and contour is more easily remembered by adults learning a new melody than the precise intervals.

The sequence of durations in a set of tones gives rise to *rhythm, tempo* (the pace or speed of the piece, loosely related to the temporal interval at which one would tap a foot or snap fingers), and *meter* (the way in which tones are perceived to be temporally grouped or organized, the most common in Western music being groups of two, three, or four). Our brains organize these fundamental perceptual attributes into higher level concepts—just as a painter arranges lines into shapes, contours, and forms. In music, these higher level concepts include melody and harmony. When we listen to music, we actually perceive multiple attributes or "dimensions" interacting.

Melodies are defined by the *pattern* or *relation* of successive pitches across time; most people have little trouble recognizing a melody that has been transposed in pitch. In fact, many melodies do not have a "correct" pitch, they just float freely in pitch space, starting anywhere one wants them to. "Happy Birthday" is an example of this, typically sung with naïve disregard to whether it is being sung in the same key from one occasion to another.

One way to think about a melody is that it is an abstract prototype, derived from specific instantiations of key, tempo, instrumentation, and so on. A melody is an auditory object that maintains its identity under certain transformations, just as a chair maintains its identity under certain transformations, such as moving it to the other side of the room, turning it upside down, or painting it red. (It was this property of melodies—the fact that their identity is defined in *relational* rather than *absolute* terms—that influenced the formation of the Gestalt psychology movement more than a hundred years ago by von Ehrenfels, Wertheimer, Koffka, and Köhler.) So for example, if you hear a song played louder than you're accustomed to, you can still identify it. If you hear it at a different tempo, played by a different instrument, or coming from a different location in space, it is still the same melody. Of course, extreme changes in any of these dimensions will render it unrecognizable; pitches outside the range of human hearing, a tempo of one beat per hour, or a loudness of 200 A-weighted decibels, dB(A), might stretch the limits of identification.

The ability to recognize transposed melodies (transformations of each pitch by equal intervals) may be innate—it has been observed in some animals and in human infants—although the neurological basis for this is not clear. The human auditory system contains frequency-sensitive neurons throughout every stage, from the cochlea on up through the inferior colliculus and the primary auditory cortex. The *absolute* pitches of tones are thus "known" to the brain from the earliest stages of processing. Yet, according to music theorist Eugene Narmour, music listening requires simultaneously attending to both the absolute and the relative pitch information.

At some stage of neural processing, a pattern must be extracted. Much of music appreciation is based on the pitting of absolute against relative cues. Consider the opening phrase of Beethoven's Fifth Symphony: three tones at the same pitch followed by a descending major third; then three more tones, all the same, at a new pitch followed by a descending minor third. Our appreciation for the piece is largely because we hear the second four tones as conceptually related to the first four—a kind of variation or extension of a theme. The fact that we hear the second group of four as similar, even though the interval is different (a minor third versus a major third), owes to a property of scales—the tones fall within the same interval class but with a different tonality.

A scale is simply a discrete set of tones used within a particular musical style or culture. Because the audible range of musical frequencies is continuous, there is no physical or objective reason why one tone should be considered the preferred basis of a musical system over any other. Indeed, the present Western standard, which sets the tone A to 440 hertz (Hz), is arbitrary and of recent origin. Just 100 years ago, European orchestras tuned to a different standard from this, and from one another—musicians traveling between cities found no consistency in tuning systems. Once one has fixed a standard tone, however, certain regularities manifest themselves across all known musical cultures. One is the octave, a frequency ratio of 2:1, which most listeners regard as sounding the most "consonant" or "pleasing" of intervals except for the unison. The perceptual relationship between tones bearing this relationship is reflected in our naming scheme: the tones at 110, 220, 440, and 880 Hz are all called "A" for example. All known musical systems have the octave and even many animal species show octave equivalence. All musical systems then divide the octave up into a set of discrete tones, usually 5 to 15, known as the scale of that culture—these are the tones that are "legal" within that system. Although variable pitched instruments, such as the violin, trombone, and voice, can produce pitches other than scale tones, the scale tones form the basis for composition and performance, and pitches outside the scale are used only for musical and expressive effects. The scale is so overlearned and well instantiated that composers can create excitement or novelty by violating scale conventions and modulating to other keys. By the age of five, human infants have learned the conventions of the music of their culture and can readily detect violations in tuning or scale.

Chords are created when three or more tones are sounded either simultaneously or in close temporal proximity. Although a melody exists independently of chords, different chords can color the melody, yielding different emotional qualities. Antonio Carlos Jobim's "One Note Samba" is an example of a tune in which the melody largely plays out on one tone, but the emotional and expressive nuances are created by the intervallic relationships between that tone and the underlying chords, or "harmony." The chord progression for a song is called its harmony. The same term, somewhat confusingly, is commonly used for a secondary musical part that accompanies the main melody, but at different pitches—two voices singing "in harmony" or two or more instruments playing together, as in a string quartet. The connection between this and the more formal, first definition of harmony is that by virtue of playing two tones at the same time (an intervallic relationship), the musicians create a "harmonic context," that is, a set of tonal relations that, according to the conventions of particular musical style, convey tension and release, according to a hierarchy of tonal stability present in the scale. Generally speaking, some scale tones are perceived as less stable (tension) than others (release), and the recognition of this hierarchy is learned at a young age through passive exposure to music, even by those without explicit musical training. The tension-release mechanism is thought to be a major underlying factor in engendering emotional responses to music.

Musical Structure, Grammar, Syntax, Semantics

Each human culture develops its own traditions for the ways in which the six perceptual attributes are employed to create music. The system of rules or conventions by which sounds are strung together in a given culture can be thought of as the grammar for that music and as reflecting a musical style, syntax, and idiom. Musical phrases are composed of notes and/or chords, but as in language, these are not randomly ordered, and a reordering of elements produces a different melody.

In probe-tone studies first introduced by Carol Krumhansl, individuals listen to a fragment of a musical sequence and are asked to judge the goodness of fit of a test or "probe" tone. Even music listeners lacking formal training demonstrate sensitivity to the structural norms of Western tonal music, assigning the highest goodness ratings to the tonic of a key, even if they lack declarative knowledge about what a tonic or a key is. This demonstrates that the average listener has implicitly internalized the rules of Western tonal music, and with increasing training, the instantiation of tonal hierarchies becomes stronger.

Experiments that introduce stimuli violating the rules of musical grammar have been employed to investigate how the human brain processes musical structure. Evidence indicates that syntax, or grammar, in language and music share a common set of circuits instantiated in frontal brain regions. Frontal brain regions have been implicated in the processing of harmonic structure and, in particular, the processing of harmonic anomalies. Several neuroimaging studies colocate musical and linguistic operations. When musical structure is disrupted, areas of the brain implicated in disruptions of linguistic syntax, Brodmann Area (BA) 47, Broca's area, and the adjoining anterior insula, are activated. The auditory cortex and the hippocampus and limbic system are activated during normal processing of both music and language. Yet music and language also employ distinct neural circuits, as evidenced by double dissociations; in general, these occur when damage to a brain region X is associated with impairment in function Y (but no impairment in function Z), while damage to a complementary region W causes the complementary pattern of performance (impairment in function Z but not Y). Following organic brain trauma, exclusive loss of musical function has been observed with little or no loss of linguistic function and vice versa.

An ongoing debate concerns the evolutionary origins of music and language, which came first, and the extent to which they coevolved. Evidence brought to bear on these questions comes from archaeological findings (for example, bone flutes at ancient human burial sites), anthropology (in particular, the study of contemporary preliterate and preindustrial societies), biology (especially the study of communication among closely related species, such as chimpanzees), and neuroscience (differential activation of brain circuits by music and language, with music tending to activate phylogenetically older structures). Some believe that music preceded language, some the reverse, but such conclusions are necessarily speculative, inductive, and not deductive. Philosopher Daniel Dennett argues that regardless of which came first, music or language, the brain no doubt had a period of coevolution over tens of thousands of years. The current neurobiological state of the human brain, proposes neuroscientist Ani Patel, is such that the syntactic processing of music and language rely on shared neural substrates, through his shared syntactic integration resource hypothesis (SSIRH).

Memory for Music and Musical Imagery

Many people report being able to hear music in their imaginations, or to play back musical selections in their mind, abilities that are a form of "auditory imagery." Experiments have shown that the average person can manipulate musical components independently through imagery, such as speeding up the tempo of a song without altering pitch or transposing the pitch without altering tempo. Timbral alterations are also relatively easy to imagine—one can imagine a well-known song played with different instrumentation, for example. For reasons that are not well understood, however, manipulating the imagined loudness of a piece of music seems to be difficult or nearly impossible for most people. When musical imagery occurs spontaneously, and without explicit intent, during the state between wakefulness and sleep, or between sleep and wakefulness, it is termed *hypnagogic* or *hypnopompic* imagery, respectively; such imagery is frequently reported to be especially vivid, detailed, and nuanced. Another manifestation of musical imagery is when music persists in the mind, known as *ear worms,* from the German word *Ohrwurm*. This phenomenon is more commonly reported in musicians as well as individuals with obsessive-compulsive disorder, and typically occurs for fragments of songs within the capacity of auditory short-term (echoic) memory, which lasts about 15 to 30 seconds.

When asked to produce songs from memory by singing, humming, or whistling, a significant proportion people produce the song at or very near the correct pitch and tempo. When there is no correct pitch or tempo (such as with folk songs or "Happy Birthday"), people tend to produce the song at a consistent pitch and tempo across multiple occasions. Even nonmusicians possess this ability, and their memory representations tend to include fine-grained features of the original performance, such as vocal nuances and performance-specific expressive details.

Imagining music has been shown to activate neural regions that are surprisingly similar to those activated when actually *listening* to music. This finding provides important support for the current theory that the process of remembering a perceptual event requires reactivation of those neural circuits that were involved in the original perception of that event.

Music can serve as an important marker for events over the course of a lifetime, and people who listen to music often tend to describe having a "soundtrack of their life." The neurocognitive basis for autobiographical music memory has only recently been studied, and activity in the rostromedial prefrontal cortex is correlated with increased memory salience for music. This same region is activated when an individual tracks a musical event through tonal space (the scale structures that represent the identity of music). Tonal space can be seen as being a precise recognition marker for a piece of music, a process necessarily involved in musical memory.

Timbre imagery has been found to induce activation in secondary auditory areas. The fact that imagery engages regions normally associated with sensory activity suggests that imagery involves mimicking the activity itself; that is, thinking about hearing involves a form of hearing itself. One of the most valued aspects of the human mind is its capacity for imagination. Musical imagery is the act of representing music in the mind without external input. Curiously, brain areas responsible for musical performance become activated when solely imagining that performance. For example, a cellist imagining he or she is playing will in fact be engaging the motor areas of the brain recruited to move the fingers as in a real performance.

Ability, Disability, Genetics, "Talent," and Musicianship

Music ability is popularly regarded to be innate—one is either born with musical "talent" or not. Part of the difficulty in distinguishing "nature" from "nurture" with music is that the child raised in a musical household—regardless of his genotype—is almost certainly apt to receive more musical input, feedback, and encouragement than the child raised in a nonmusical household. In one study of conservatory students, the amount of practice over the four years of their instruction was a far greater predictor of final year ability than were the ratings of potential given to them on intake. If by "talent" one means a set of unobservable, hard-wired propensities that are more predictive than factors such as hard work and time-on-task, there exists no credible evidence for it. Self-reports of world-class musicians, as well as experimental studies, point strongly to the view that musicians are made not born and that practice accounts for an overwhelmingly large proportion of the variance in who becomes an expert musician and who doesn't. Those factors that cause some musicians to practice more than others may well have a genetic component, attributes such as single-mindedness, seriousness, and conscientiousness.

There may also exist genetic components contributing to attention span, auditory memory, and auditory sensitivity, qualities that would be important for success in music, but not exclusively so. It is unlikely that there exist a "music gene" or alleles that determine musicality, but there may well be genetic contributions to many of the component skills that are necessary to become a great musician, such as those cognitive factors previously listed, as well as physical factors, such as reflexes, finger speed, or motor dexterity.

Confounding efforts to study musical ability is the fact that it can manifest itself in many forms. One can be exquisitely sensitive to music without even being a musician—many listeners display intense emotional reactions to music without the ability to play it, and disc jockeys, film music supervisors, and record company talent scouts lacking formal musical training are not necessarily at a disadvantage. One can be an expert in one domain of music and lack ability in another. There are composers who lack instrumental ability (e.g., Irving Berlin, who could barely play his own compositions on the piano), instrumentalists and conductors who do not compose (Artur Rubinstein, Herbert von Karajan), and great musicians may read music or not, and improvise or not.

A small percentage of the population appears to lack musical ability or sensitivity. When present from birth, this has been labeled *congenital amusia;* when it is the result of organic brain disease or trauma, it is called *acquired amusia.* In the popular press, these terms are used interchangeably with the terms *tone deafness* and *tin-ear syndrome.* However, the amusias no doubt comprise a heterogeneous set of disabilities with distinct etiologies. There are individuals who cannot identify songs and others who have identification defects but can't sing in tune, producing abnormal variability in the tones they generate. Specific deficits in rhythm, pitch, and timbre have also been observed, both following brain injury and congenitally. Dissociations between musical alexia (inability to read) and agraphia (inability to write) have also been reported.

The study of distinct, well-defined, and atypical populations is important because it offers a unique opportunity to investigate specific aspects of cognition and to establish the degree to which various cognitive abilities are correlated with, or can be decoupled from, one another. In particular, the study of populations with genotypic abnormalities (including Williams syndrome, autism spectrum disorders, and Down syndrome) have sparked new debates regarding the modularity of brain function, independence of mental faculties, and theories of neural organization. Members of all three groups tend to be unusually attracted to music, but what they are hearing and perceiving remains open to debate. Individuals with Williams syndrome, characterized by low IQ and deficits in spatial ability, reasoning, and numeracy, show relatively intact language and musical abilities, yet are often just as attracted to odd sounds (such as motors, fans, and water running) as they are to symphonies and pop songs. Individuals with autism spectrum disorders, in spite of an attraction to music and a sometimes eidetic memory for sound ("phonographic memory") appear to be insensitive to many of the nuances in musical emotion.

Another rare but unique aspect of pitch processing is some people's capacity for absolute pitch. Possessors of absolute pitch can name a note when it is heard in isolation, provided they have knowledge of the labels assigned to musical notes (A-G, or do-re-mi); there are also ways of identifying absolute pitch experimentally in the absence of labels. Absolute pitch possessors do not require the anchor of another note for pitch identification by means of calculating musical intervals (relative pitch). Some musicians have acquired a strong representation for an isolated musical tone (usually the tone to which they are accustomed to tuning their instrument). They use this memorized tone to calculate other notes using relative pitch, and this phenomenon is called quasi-absolute pitch.

The benefits of musicianship have long been hailed for their potential effects on cognition. Childhood music lessons have a small but long-lasting positive correlation with IQ and with academic performance. Recent studies have provided the first concrete evidence that playing a musical instrument significantly enhances the brain stem's sensitivity to speech sounds. Preliminary work on rhythmic training suggests that it may help to ameliorate dyslexia. Correlational studies further suggest that a mastery of instrumental music may lead to a positive self-concept, but carryover effects to other domains have not yet been demonstrated. Brain imaging studies suggest that early training has its greatest effect on neural systems involved in sensorimotor integration and timing.

Emotion and Expectation

Music represents a dynamic form of emotion, and the conveying of emotion is considered to be the essence, if not the purpose, of music, and the reason that most people report spending large amounts of time listening to music. Recently, much study has been focused on the biological underpinnings of musical emotion, particularly the involvement of neural reward systems. This phenomenon has been studied through investigating the chill response, a physical sensation up the spine that can occur as a result of conscious music listening. It varies from individual to individual, and is based on a number of factors, such as structural components and loudness of the music, as well as character/personality organization and musical experience. When people listen to music that they report consistently gives them chills, blood flow increases to those centers of that brain that are implicated in reward, emotion, and arousal: the amygdala, nucleus accumbens, and ventral tegmental area. These regions modulate levels of dopamine. In addition, activation is observed in the hypothalamus, insula, and cerebellar vermis.

The experience of pleasant, or consonant, music also activates orbitofrontal, subcallosal cingulate, and frontal polar cortical areas. The hippocampus has been found in positron emission tomography studies to activate during pleasant music, and the parahippocampal gyrus, also implicated in emotion processing, has been found to activate during unpleasant, or dissonant, music.

Much of our emotional reactions to music are believed to be caused by the meeting and violating of musical expectations. Listeners track the progression of music over time, noting the pitches and rhythms employed, and form subconscious predictions about what will occur next. A musical piece that we find pleasing strikes the balance between meeting those predictions some of the time and violating them in interesting ways the rest of the time. These predictions may involve statistical maps of which notes are most likely to follow certain melodic and harmonic progressions, and the tracking of such tonal movement is now thought to involve regions in BA47, BA44, and the rostral portion of the ventromedial superior frontal gyrus and the right orbitofrontal gyrus.

There exists a widespread belief in Western culture that major keys are associated with positive affect, or happiness, while minor keys are related to negative affect, or sadness. In fact, this occurrence turns out to be largely a product of musical exposure and learning and is thus culturally dependent. It has been shown that other musical systems (e.g., Middle Eastern, Indian) do not share these associations.

Hemispheric Specialization

Early reports stated that music is predominantly a right-hemisphere activity and language a left-hemisphere activity (in neurologically intact right-handed listeners). This is now considered to be an oversimplification, in part because of the distributed nature of specialized processing mechanisms acting on the individual musical attributes previously listed. It is now known that music listening, performing, and composing engage regions throughout the brain, bilaterally, and in the cortex, neocortex, paleo- and neocerebellum. Laterality effects do exist, however. For example, magnetoencephalography (MEG) responses to deviations in the memorized lyrics of tunes are stronger in the left hemisphere, whereas the perception of violations of expected notes are governed by the right hemisphere. The act of learning music causes a left hemisphere shift as naming processes become involved.

Pitch and rhythm have been found to be neurally separable. The processing of melodic, but not rhythmic information of music is neurally isolated in the auditory association cortex of the superior temporal gyrus, as documented by lesioned patients. Based on evidence from brain-damaged patients, Isabelle Peretz and Max Coltheart suggest a theoretical model of functional architecture whereby music processing modules, grouped into pitch organization and temporal organization, represent an interactive and separable system of music processing. Timbre appears to invoke distinct neural circuitry as well, bilaterally in the temporal lobes. Recent evidence suggests that timbre maps, similar to tonotopic pitch maps, may exist in the cortex.

Music Preferences and Individual Differences

Measures of personality and individual differences have been shown to correlate with taste in music; the correlations are relatively small, but both significant and robust. Using the Big Five personality inventory and a cross-section of songs representing the major genres and subgenres of Western tonal music, certain consistencies have emerged. Although such research is still in its early stages, extraverted individuals tend to like music that is characterized as energetic and rhythmic. Individuals who rate high on factor 5, openness to new experience, show no correlation with such music, but rather with music that is described as reflective and complex. Upbeat and conventional music correlates with factor 2, agreeableness.

Several hypotheses exist as to why musical taste might be related to personality. In some cases, people may prefer and seek out styles of music that reflect and reinforce aspects of their personalities. Research suggests that personality influences how individuals think, feel, and behave. For example, that people with high levels of extraversion seek out situations that allow them to be talkative and sociable. In contrast, more introverted people tend to seek out environments where they have limited contact with other people, especially people they don't know. Just as people seek out and create social environments that reinforce aspects of their personalities, so too might people seek out auditory or musical environments that reinforce aspects of their personalities. Those who are normally extraverted, for example, may help to maintain their self-identity and energy level by listening to energetic music. Because music is a part of social identity in contemporary society, people may also seek out music that they believe will create a desirable impression of them. Adolescents, in particular, use music as a badge to communicate their status and affiliation with a particular peer group. Finally, many individuals report using music for mood induction. This further underscores the role of music in emotional regulation, and its role in maintaining fundamental aspects of well-being.

The study of music perception and cognition has informed fundamental issues in cognitive psychology and cognitive neuroscience, providing a window into higher cognitive function. As a universal human activity, music's evolutionary roots are no doubt very old, and the development of the brain, mind, and culture would have occurred alongside changes in the way we create, perform, listen to, and use music in everyday life. Although far more experimental psychologists study visual processes than auditory processes, the nature of mental representations for music—their vividness, relatively early formation in infants, and durability—allows for the scientific study of cognition from an alternative perspective, using ecologically valid stimuli that are meaningful to many human experimental participants.

Daniel J. Levitin and Anna K. Tirovolas

See also Auditory Scene Analysis; Emotional Influences on Perception; Melody Perception; Sound Reproduction and Perception; Timbre Perception

Further Readings

Deutsch, D. (Ed.). (1999). *The psychology of music* (2nd ed.). San Diego, CA: Academic Press.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation.* Cambridge: MIT Press.

Juslin, P., & Sloboda, J. A. (2001). *Music and emotion: Theory and research.* Oxford, UK: Oxford University Press.

Levitin, D. J. (2008). *The world in six songs: How the musical brain created human nature.* New York: Penguin.

Meyer, L. (1956). *Emotion and meaning in music.* Chicago: University of Chicago Press.

Mithen, S. (2006). *The singing Neanderthals: The origins of music, language, mind and body.* Cambridge, MA: Harvard University Press.

Patel, A. D. (2007). *Music, language, and the brain.* New York: Oxford University Press.

MUSIC IN FILM

Music and film are both based on communication of meanings through time. The emotional, mood, and perceptual effects of music in film can reinforce visual and textual components or act independently of the dramatic elements as a counterpoint to dramatic visual/textual elements, and thus create satire or irony. The study of these effects involves measuring subject response to combinations of music and film and determining goodness of fit, changes in affect or emotional response, and changes in attention or memory. Measurement has often been achieved by rating scales, based on variations of the semantic differential, an approach using bipolar opposite words in three factors, such as good-bad (evaluative), active-passive (activity), and strong-weak (potency), thus operationalizing dependent variables, such as affect, meaning, and emotion. Independent variables have included the use of visual elements alone, musical elements alone, and their combination for comparative analysis. This entry considers research on the interaction between music and visual elements of film.

Sandra Marshall and Annabel Cohen conducted an early series of experiments on how music affects subject characterization of animation figures relative to musical combinations. Animations of a small triangle, large triangle, and small circle were combined with strong and weak composed music prototypes. The weak music, in C major with high tessitura (pitch/frequency range), was contrasted with strong music, in Aeolian minor in low tessitura. Subject ratings on evaluative, potency, and activity scales were collected for the animation alone, the music alone, and combinations thereof. Results indicated that the large triangle had the highest activity and potency ratings and the lowest (most unpleasant) evaluative ratings. They found that ratings of the animation alone were altered by combination with the music, particularly as the small triangle increased in activity with strong music. Based on these results, Marshall and Cohen proposed the congruence-associationist model (CAM). Congruence is the consonant (appropriate) connection of ratings of the visual and musical elements. The proposed model of association (e.g., big triangle = strong, fast and minor music = strong) and time congruence, loosely connected to tempo and accent, led to a greatly expanded model by Cohen that includes text, speech, visual surface, music surface, and sound special effects, elements in relation to short-term and long-term memory. Most of these elements remain unexplored in experimental research as of this writing.

Scott Lipscomb and Roger Kendall conducted an experiment focused on whether the composer's intent in combining music and film was communicated to the perceiver. Five excerpts from the film *Star Trek IV* (music by Leonard Rosenman) were presented in their original combination of music and visuals and in all crossed combinations. Subjects were asked to rate the degree of fit in randomly presented excerpts. The combinations were carefully edited by film composition students for temporal congruency, defined as the coordination of music and visual accent structures (periodic times of attention), a variable of perceptual importance. They found that the composer-intended excerpts were rated as best fit overall. Excerpts of visual and musical elements that were less concrete in their use of time organization and musical elements, such as musical tonality and repetitive accent (beat) and visually abstract patterns, were most often mismatched with the composer intent.

Lipscomb and Kendall's model of perception involved the interaction in combination of purely associative musical elements (such as major = good, minor = bad) as well as the effects of time alignment of visual accents (points of attention) with musical accents (pulse/beat). Lipscomb conducted further research with film and animation visuals alone and in combination with varying musical elements, and found that subjects could discern a visual accent structure (pulse) and musical accent structure, both measured by subjects tapping a computer key, in isolation. When combined in musical/visual congruence, subject ratings on goodness of fit and semantic differential scales were as hypothesized; incongruent alignments often changed the goodness of fit and semantic ratings of the composites.

Kendall hypothesized that aspects of semiotic theory in cognitive musicology, as outlined by W. J. Dowling and Dane Harwood, could provide a useful approach to understanding film, animation, dance, and other temporally organized visual forms in relation to music variables. He combined the semiotic theories of Charles Peirce and Leonard Meyer in proposing a continuum of referentiality. At one end of the continuum, the meaning of musical elements is arbitrary in connection to the visual; an example is the use of a national anthem, such as *La Marseillaise* in *Casablanca,* with the positive characters and actions of the French underground. Like the American national anthem, the association of the country with the music is entirely arbitrary, such as the connection of the symbol "cat" with the animal. The extrareferential (pointing outside the scene) elements in film, as previously noted, are contrasted with intrareferential elements, such as thematic material generated internally within the film; this is the idea of the *leitmotiv,* extensively used in *Gone with the Wind,* for example, as in the Tara theme. At the other end of the continuum of referentiality are connections that are purely syntactical, the alignment in time of accent structures in the visual and musical domains, prevalent in dance, ice skating, and abstract animation, extensively seen in parts of the film *Fantasia.* In the center of this continuum are meanings, which are iconic, where the pattern of visual and/or musical elements suggests connections to one another (a visual example is the weeping willow, so named because the branches of the tree descend in arches to the ground). In film, a cartoon character falling from a cliff has descending visual motion accompanied by descending pitch. Combinations of these elements, association, iconicity, and pure temporal congruence are found in the *E.T.* bicycle sequence toward the end of the film. Visual accents and musical accents connect the pedaling together in time, the *E.T.* theme is associated (intrareferentially) with the character, and patterns of arches in pitch follow the movement of the bikes over the hills and as they rise in the flying sequence. Using a continuous rating scale from referential (association) to iconic to pure accent, musically trained subjects could reliably rate magnitudes of relations in musical/visual composites drawn from film, TV, dance, and ice skating involving varying magnitudes of association, iconicity, and pure accent alignment, suggesting noncategorical judgment.

In a series of studies by Marilyn Boltz, central aspects of congruence in mood association were systematically studied in recall of dramatic intent, musical recall, visual recall, and discrimination. Subjects could recall mood-congruent combinations of film and music better than noncongruent combinations (such as major = positive and minor = negative). In incongruent combinations, the music and film elements were separately encoded, with results indicating conflicts between these dimensions when combined. Results suggest a combination of film and music elements in mood-congruent contexts are stored in memory, whereas separate attention is paid to film and music elements that are disparate in associative mood. The use of incongruent combinations is well-explored in film; for example, in *Clockwork Orange,* some violent scenes, such as the rape scene, are accompanied by music that is positive in mood (*Singin' in the Rain*). The incongruence results in irony and satire, fundamental themes of the movie.

In these and other studies, the connection of timing accents (temporal congruence) and association expectations (mood congruence) leads to better memory encoding, goodness of fit ratings, and higher evaluative ratings. Denotative meanings in arbitrary mood association (major = good, minor = bad, fast = active, slow = passive, for example) are complemented by temporal/mood associations that are connotative and iconic. Caution in this area is also suggested, in that common expectations of congruence can be intentionally violated by the artists to draw particular attention to the visual dramatic action with great effect. Temporally dynamic unfolding of meanings in the visual, textual-dramatic, and musical domains has only limited scientific study; new methodologies need to explore how such meanings unfold through long periods of time within the film. This problem is no different from that in many perceptual domains, where single ratings are taken after a small or moderate length example. There is much yet to be explored in this new multimedia world.

Roger A. Kendall

See also Emotional Influences on Perception; Film (Cinema) Perception; Multimodal Interactions: Visual–Auditory; Music Cognition and Perception

Further Readings

Boltz, M. (2005). The cognitive processing of film and musical soundtracks. *Memory and Cognition, 32,* 1194–1205.

Cohen, A. (Ed.). (1994, 1996). The psychology of film music (Special two-issue volume). *Psychomusicology, 13.*

Kendall, R. A., & Savage, W. H. (Eds.). (2005). Perspectives in systematic musicology. *Selected Reports in Ethnomusicology* (Vol. 12, chaps. 1–3). Los Angeles: Department of Ethnomusicology Press UCLA.

N

NAÏVE REALISM

Naïve realism is the name for a philosophical theory of visual perception and the nature of visual experiences (though the theory may be generalized to the other senses). According to naïve realism, veridical visual experiences are relations to material objects that exist whether or not they are perceived and so are mind-independent (naïve realism is therefore sometimes referred to as the *relational* theory of perception); such experiences have material objects as constituents—the objects of visual awareness are literally part of the experience. Because hallucinatory experiences are not experiences of material objects, it follows that you could not be having the visual experience that you are now having if you were having a hallucination. Therefore, naïve realists hold a *disjunctive* theory of visual experiences: A visual experience in which it seems to you that there is a book on the desk in front of you is *either* a psychological episode that actually has the book as a constituent *or* a fundamentally different kind of psychological episode that merely seems to have a book as a constituent.

Why should we accept naïve realism? One reason is that it captures how our visual experience seems to be to us in introspection. When you look at a book lying on the desk in front of you, the book visually appears to you as being a certain way: as having a certain size and shape, for example, and as being a certain color. Attending to the book and its features is a way of finding out how the book—the object of your experience—appears. Suppose now that you introspectively reflect on your visual experience of the book: You shift your attention from the book to your experience of the book. Shifting your attention from the book to your experience of the book is not shifting your attention from one object to another—in introspectively reflecting on your experience of the book there is no other object for you to attend to than the book itself (this is sometimes referred to as the *transparency* of visual experience: Visual experience seems transparent to the world). Your experience introspectively seems a certain way to you. In characterizing how your *experience*—a psychological episode—seems, you can do no better than describe how the *objects* of your experience—material objects—appear. Your experience seems to present the book and its features, and the features that determine the character of your experience are features that seem to be instantiated by the book. Therefore, your visual experience seems to involve the presentation of a material object—the book—and its features, and any account of visual experience that rejects this is committed to claiming that visual experience is other than it seems to be introspectively. A second reason for accepting naïve realism is epistemological. If the visual experience you enjoy when seeing a book is of a kind that could occur in the absence of the book (as it would do, for example, when you have a hallucination of the book), then it does not provide you with the kind of cognitive contact with the world required to ground knowledge. Therefore, the experience you have when you see the book must be of a kind that could not occur in the absence of the book.

Objections to naïve realism usually focus on the problem of explaining hallucinatory experiences. For any veridical experience of a material object that a person has, it is possible to bring about—by, one might suppose, direct stimulation of one's brain—an exactly similar hallucinatory experience—an experience that the person cannot introspectively tell apart from the veridical experience. Although such a hallucinatory experience is introspectively exactly similar to the veridical experience, it does not involve a relation to any material object. Therefore, it must be possible to explain the introspectible character of the hallucinatory experience (where introspectible is the character of the hallucinatory experience discoverable by introspection) in a way that does not involve any relation to material objects. If it is possible to explain the introspectible character of hallucinatory experiences without appealing to material objects, then it must be possible to explain the introspectible character of any exactly similar veridical experience without appealing to material objects. Therefore, it is possible to explain the introspectible character of any visual experience in a way that does not involve material objects, and that undermines the claim that naïve realism explains how our veridical visual experience introspectively seems. In order to respond to this objection the naïve realist must deny that it is possible to explain the introspectible character of hallucinatory experiences in a way that doesn't involve any relation to material objects.

Naïve realism is a form of direct realism about perception. It is a form of *realism* because it supposes that the objects of our visual awareness are mind-independent material objects. It is a form of *direct* perception because it takes our awareness of objects to be direct. We are aware of an object directly if we are not aware of it indirectly, and we are aware of an object indirectly if we are aware of it in virtue of being aware of something else. Indirect theories of perception have typically claimed that we are aware of mind-independent material objects only in virtue of being aware of mind-dependent objects or *sense data.* Naïve realism is not the only direct realist theory of perception. According to the intentional or representational theory of perception, perceptual experiences are like beliefs in having intentional or representational content. A visual experience represents mind-independent material objects as being a certain way and is veridical if and only if those objects are the way they are represented to be. The objects of our visual awareness are the mind-independent material objects that are represented by our visual experience: We are aware of them in virtue of having an experience that veridically represents them, and not in virtue of being aware of something else. Naïve realism denies that the objects of our awareness are the objects that are veridically represented by our visual experience: According to naïve realism, experience does not *represent* its objects, it has its objects as constituents.

Matthew Nudds

See also Indirect Nature of Perception; Intentionality and Perception; Object Perception; Perceptual Representation (Philosophy); Philosophical Approaches

Further Readings

Hinton, J. M. (1973). *Experiences: An enquiry into some ambiguities.* Oxford, UK: Clarendon Press.

Martin, M. G. F. (2002). The transparency of experience. *Mind and Language, 17,* 375–425.

Martin, M. G. F. (2006). On being alienated. In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 354–410). Oxford, UK: Oxford University Press.

McDowell, J. (1982). Criteria, defeasibility and knowledge. *Proceedings of the British Academy, 68,* 455–479.

Travis, Charles. (2004). The silence of the senses. *Mind, 113,* 57–94.

NATURE AND NURTURE IN PERCEPTION

Probably no other issue has engendered more debate in philosophical and scientific discourse than the nature–nurture dichotomy. This entry describes some of the arguments for each side of the dichotomy and the evidence that developmental outcomes are, in fact, the *joint* product of the coaction of interdependent processes with neither having priority over the other. The idea of separate processes first arose when Plato and Aristotle asked the ultimate epistemological question: What is the origin of human knowledge? In essence, these two philosophers wondered where our ideas of physical objects, causality, number, and space come from. Plato's answer was that our senses do not provide adequate information to form abstract ideas about our world and his "poverty of the stimulus" argument forced him to conclude that all knowledge was innate. Aristotle rejected Plato's argument and declared that our senses do, in fact, provide all the necessary information to understand our world and, thus, concluded that knowledge is acquired through experience. Centuries later, the rationalist philosophers Rene Descartes and Immanuel Kant followed in Plato's footsteps and also argued that we are endowed with a set of innate abstract ideas (i.e., concepts and principles). For the rationalists, perception also was too impoverished to provide enough information about the world and, thus, they held that we must rely on our abstract ideas together with our faculty of reason to arrive at an understanding of our world. John Locke and David Hume, the British empiricist philosophers, challenged the rationalist argument and, like Aristotle, concluded that perception is sufficient to specify our world. Consequently, Locke and Hume asserted that all human knowledge is acquired through experience.

The philosophical arguments regarding the origins of knowledge are inexorably tied up with questions regarding the nature of individual development and the roots of organic form. Aristotle understood this link and, as a result, studied the development of chicks. He noted that development is initiated by some vital force acting on unformed organic matter. As development progresses, this organic matter becomes gradually differentiated into a complex organism. The process of developmental transformation became known as *epigenesis*. A contrasting view that was prevalent at that time, and one that was held for many centuries after Aristotle, was *preformationism*. According to this view, the egg or sperm contained a fully formed adult version of the human being (the homunculus) in miniature form and development consists of its gradual growth into a fully mature organism. Although it is evident today that preformationism is wrong, it was not until the advent of the microscope and the emergence of the field of embryology that this view was discredited. *Predeterminism*—the notion that organisms pass through qualitatively different stages of organization—replaced preformationism. This was a more nuanced conceptualization of development in that it acknowledged the epigenetic principle of transformation but, nonetheless, retained the notion that all knowledge is innate and present at birth.

Charles Darwin's publication of the theory of evolution in 1859 had important but, it could be argued, contradictory effects on developmental thinking. On one hand, the core assumption underlying evolutionary thinking is that the evolution of organisms—*phylogenesis*—is driven by a process of transformation. This is the same process that underlies the process of individual development—*ontogenesis*—and, thus, it would be reasonable to expect that this core assumption might have brought the concept of epigenesis to the fore. It did not, though, because the other core assumption underlying the theory of evolution was that natural selection operates on inherited, species-specific characteristics. Thus, it was natural to conclude that ontogenesis begins with an unformed but fully predetermined organism, and this lent further credence to Plato's, Descartes', and Kant's nativist arguments that human nature and knowledge could be passed down through the generations in the form of inherited characteristics. When this idea was later combined with Gregor Mendel's experimental proof of inheritance, the linking of inheritance with phenotypes (i.e., organic form) by the modern synthesis, and James Watson and Francis Crick's discovery of DNA as the physical basis of inheritance, the stage was set for equating the concept of "nature" with the concept of "genetic." Thus was born the modern version of the nativist idea that the knowledge that is given at birth is coded in our genetic blueprint, and that the sensory, perceptual, and cognitive skills that emerge in development represent an unfolding of that blueprint. Moreover, this modern nativist idea was combined with the concept of genetic encapsulation—the idea that genes are not subject to external influences—and together this conceptualization of development began to dominate biological and psychological thinking.

The problem with the claim that knowledge is coded in our genes and that it unfolds during development is that it overlooks the basic biological fact that genes only code for proteins. In addition, many events that are external to the organism's genetic endowment play a critical role in the organism's development; without them the organism either does not develop or develops in completely different ways. This is obvious from experiments showing that animals with identical genetic makeup develop into dramatically different organisms when such seemingly mundane factors as temperature or chemical environment are manipulated. In other words, genes are not encapsulated entities but, rather, work hand-in-hand with their environment.

Shifting Conceptualizations

The nonobvious effects of exogenous factors, such as temperature and chemical context, on developmental outcome began to permeate the thinking of developmental psychobiologists starting in the late 1950s and led them to initiate studies to investigate the influence of stimulus context and experience on development. Importantly, experience was construed broadly and included not only traditional learning processes but all forms of external and internal (including self-produced) stimulation. These studies made it clear that conceptualizing development in terms of the nature–nurture dichotomy was too simplistic, and that many factors besides the traditional process of learning contribute to the developmental emergence of behavioral capacities. For example, these studies showed that rearing conditions, stimulation history, hormonal/neural status, and social factors all play important and sometimes nonobvious roles in development. Consequently, it became obvious that partitioning the causal origins of behavioral development into nature (i.e., genes) and nurture (i.e., learning) was misleading, and that a more accurate way of thinking about development was that it is a process consisting of the bidirectional interaction between the organism (i.e., all endogenous factors) and its environment (i.e., all exogenous factors).

This new conceptual framework—developed by T. C. Schneirla, Zing-Yang Kuo, and Gilbert Gottlieb—was dubbed *probabilistic epigenesis* and has been effectively subsumed within a general approach to the study of behavioral development that eventually came to be known as *developmental systems theory* (DST). According to DST, the development of an organism and its behavioral capacities is the result of a complex interaction and coaction of different factors located at different levels of organization. These factors consist of genes that are embedded inside cells and their complex chemical environments, cells (e.g., neurons) that are embedded in organs (e.g., the hippocampus), organs that are embedded in systems (e.g., the brain), and systems that are embedded in societies. The emergence of a new structure (e.g., the central nervous system) and/or function (e.g., sensation, perception, cognition) is due to bidirectional horizontal interactions (e.g., gene-gene, cell-cell, organ-organ, system-system), as well as vertical interactions (e.g., gene-cell, gene-organ, gene-system, cell-organ, cell-system, organ-system). For example, genes can code for the production of specific hormones leading to specific reproductive behaviors in males, but the production of those hormones (mediated through gene expression) is initiated by changing light input as well as the presence of a receptive female. In other words, initiation of male reproductive behavior is not due to genes or to the availability of sensory inputs; it is due to both factors working together. Thus, according to DST, newly emerging structures and functions do not represent the unfolding of predetermined characteristics inscribed in the organism's genetic blueprint. Rather, they represent novel and emergent properties of the developmental process constructed by experience and the developmental process per se.

Most contemporary behavioral scientists readily accept the concept of interaction and apply developmental systems principles when formulating research questions and interpreting empirical findings. Nonetheless, some contemporary cognitive scientists continue to insist that we are either endowed with perceptual "primitives" or "core knowledge" at birth and that all the subsequent knowledge that is acquired through learning builds upon these primitives or core knowledge. As proof of the predetermined nature of development and of the existence of primitives or core knowledge, these scientists often cite the fact that human infants understand objects' physical properties and their behavior. Adherents of DST counter that the infants who exhibit knowledge of core physical principles are not newborns and, thus, they have already had all sorts of postnatal experience. In addition, they point out that finding that a particular skill emerges at birth, or, for that matter, at any other point in prenatal or postnatal development, is only interesting insofar as suggesting a starting point for an inquiry into the myriad and complex developmental determinants that underlie the emergence of this skill.

Role of Experience in Perceptual Development

The broad concept of experience is critical to DST and its pivotal role in perceptual development has been supported by a wealth of empirical evidence from animal and human infant studies. This evidence has provided important new insights into the developmental processes underlying the emergence of specific perceptual skills and has shown that the nature–nurture debate is no longer useful for framing developmental questions. This body of evidence comes from two types of studies. Some of them systematically assess the effects of specific stimulus conditions normally present in an infant's ecology on responsiveness and developmental outcome. Others manipulate or deprive the infant of particular sensory input(s). Although the former types of studies can be carried out in humans, the latter can only be done either in animals or in human clinical populations with specific kinds of developmental abnormalities (e.g., individuals who are deaf or who have congenital cataracts).

Animal studies have provided insights into the effects of developmental history and experience on normal perceptual functioning and evidence of remarkable early plasticity that enables developing organisms to adapt to their ecological setting. For example, birds, like all other species, must recognize their conspecifics in order to survive and successfully reproduce. Some bird species exhibit species-specific recognition at hatching, whereas others require a fairly long period of perceptual experience. When species-specific recognition appears at hatching, it is tempting to interpret its apparently sudden appearance as a manifestation of an innate and genetically determined perceptual skill. Indeed, the ethologists Konrad Lorenz and Nikolaas Tinbergen did just that and considered this to be a manifestation of instinctive behavior. However, subsequent investigations have shown that the development of species-specific recognition, regardless of whether it appears at hatching or later in postnatal life, is actually a product of the complex interaction and coaction of multiple factors. For example, newly hatched mallard ducks exhibit a preference for their mother's species-specific appearance and audible call. Studies have shown that preference for the maternal call is induced prenatally and is the result of the embryo hearing its broodmates' vocalizations and its own prehatching vocalizations; when duck eggs are incubated in isolation and/or when embryos are devocalized, they no longer prefer the maternal call. Interestingly, human newborns also exhibit a preference for their mother's voice and, like in birds, this preference is the result of exposure to their mother's voice prior to birth.

Studies of the development of song in various species of birds also have provided proof of the power of perceptual experience. When songbirds are young, they produce subsongs that differ qualitatively from adult songs in that they are highly variable in terms of structure and timing, are sung at low amplitude, and include elements not contained in the adult song. Later, young birds begin to produce what are known as plastic songs. These now include the notes and whistles found in adult songs, but they still are poorly articulated and not sung in a stable order initially and only gradually become adultlike. The process by which the birds make the transition from subsongs to plastic songs to mature songs relies heavily on feedback from adult conspecifics. This feedback consists of countersinging and of social shaping of the vocal precursors. Moreover, this process is characterized by an initial overproduction of song elements and a gradual and selective attrition of inappropriate ones. Again, the parallel between the development of bird song and the development of infant babbling in human infants is quite striking, despite the vast difference in phyletic level. Thus, like in songbirds, social shaping plays an important role in the development of canonical babbling in human infants. Those infants whose babbling sounds are selectively and contingently reinforced by the caregiver restructuring their babbling achieve a closer match between their phonological forms and those of the caregiver.

Similar effects of early stimulus history and experiences on the development of basic perceptual abilities have been reported in studies of the development of spatial localization. For example, owls locate their prey by using a combination of multisensory (i.e., visual and auditory) cues. When used together, such multisensory cues exert synergistic effects and, thus, make it easier for owls to locate and capture their prey. Studies have shown that barn owls only can take advantage of the multisensory synergistic effects if the neural representations of visual and auditory space overlap and that this overlap is sculpted by early experience. This is evident in the finding that when the correspondence between auditory and visual inputs is disrupted by prisms that shift the visual field to the right or left, the owls' behavioral responses to multisensory cues are disrupted because the neural representations of auditory and visual space no longer overlap.

Research in cats also has indicated that behavioral multisensory localization responses develop gradually. Experiments have shown that the multisensory neurons in the superior colliculus and the anterior ectosylvian cortex of the cat that control behavioral localization only gradually become functional after birth, and that postnatal experience is essential to their development. When kittens are deprived of normal postnatal visual input by being raised in the dark, their multisensory neurons do not perform normal intersensory integration.

As in owls and cats, audiovisual spatial localization skills develop gradually in human infants during the first year of postnatal life. Although newborns can localize a laterally presented sound by turning their head in its direction and looking at it at the same time, it is not until eight months of age that the latency of the spatial localization response becomes adultlike (i.e., it becomes a nonlinear combination of the auditory- and visual-only response latency). This finding suggests that the gradual development of adultlike localization behaviors is probably partly based on the experience that infants have with lateralized sights and sounds during the first months of postnatal life.

Speech and language is unique to humans and, as a result, the nature–nurture debate has been especially contentious in this domain. Nonetheless, a growing body of evidence has been slowly shifting the debate away from nativist claims that language is an innate and genetically coded skill to a more balanced view that focuses on developmental processes and the contribution of experience. This shift is partly due to the discovery in the early 1980s that, regardless of the language environment into which infants are born, they start out with a broad sensitivity to the speech sounds from all languages. However, as infants grow and acquire experience with their native language, their sensitivity to nonnative speech sounds declines in a manner similar to the selective attrition of inappropriate song elements in developing songbirds. For example, young English-learning infants can detect the English phonetic contrast between the syllables /ba/ and /da/ as well as the Hindi phonetic contrast between the dental syllable /ta/ (produced by placing the tongue against the teeth) and the retroflex syllable /Ta/ (produced by curling the tongue and placing its underside against the roof of the mouth), whereas older English-learning infants no longer detect the Hindi contrast. This decline in infant sensitivity to nonnative auditory speech is paralleled by a decline in infant sensitivity to visual speech. Whereas four- and six-month-old English-learning infants can discriminate between silent French and English visual speech articulations, eight-month-old English-learning infants no longer do. This decline is a direct result of language-specific experience because infants who grow up in bilingual homes maintain the ability to discriminate between them.

The decline in sensitivity observed in infant response to auditory and visual speech also occurs in other domains, indicating that this kind of perceptual narrowing reflects the operation of a general developmental process. For example, 6-month-old American infants can perceive Western as well as non-Western musical rhythms, whereas 12-month-old infants no longer respond to non-Western rhythms. Moreover, the decline in sensitivity to non-Western musical rhythms can be prevented in infants through additional exposure to non-Western rhythms. Interestingly, the decline cannot be prevented in adults with additional exposure to non-Western rhythms, indicating that perceptual plasticity is confined to early life.

In addition to perceptual narrowing in response to auditory and visual speech and to musical rhythms, the infant's response to silent nonvocalizing faces also declines. Thus, whereas six-month-old infants can discriminate between different human faces and between different monkey faces, nine-month-olds only discriminate between different human faces. Similarly, whereas younger infants can discriminate the faces of other races, older infants only can discriminate the faces of their own race. As in music perception, the decline in sensitivity to nonnative faces can be prevented by providing additional experience with nonnative faces during the period of decline.

Finally, as might be expected from the findings of experience-dependent perceptual narrowing of responsiveness to auditory and visual speech, music, and faces, infant responsiveness to audio-visual relations also declines during infancy. Thus, even though infants get better at perceiving the correspondence between human facial gestures and accompanying human vocalizations as they get older, this is not the case when the facial gestures and vocalizations are those of another species. For example, four- to six-month-old infants can match nonnative (i.e., monkey) faces and their vocalizations, but older infants (i.e., eight months of age and older) no longer do. When this finding of intersensory perceptual narrowing is considered together with the findings of perceptual narrowing in single sensory modalities, it becomes clear that perceptual narrowing is a domain-general and pansensory process. In addition, it becomes clear that this process is driven by early and selective experience with native sensory inputs that lead to the narrowing of an initially broad perceptual sensitivity down to one that leaves the infant with a sensitivity that best matches his or her ecological exigencies.

Moving Beyond the Dichotomy

Although the nature–nurture dichotomy has a long and distinguished intellectual history, modern science has shown that it is no longer a useful conceptual tool for posing questions about development. The empirical results reviewed here demonstrate that endogenous and exogenous factors jointly determine perceptual development and that the many components of particular perceptual skills are constructed over many different but overlapping and interacting developmental time scales. Indeed, some of the most convincing and dramatic evidence of the interdependent nature of development—even at the genetic level—comes from studies of the effects of maternal behavior styles on the development of offspring. These studies show that mice who are bred for fearfulness and high stress reactivity can be turned into docile, low-reactive animals despite their presumably genetically determined nature. This can be accomplished by simply cross-fostering their young with the mothers of a docile inbred line of mice who typically lick and groom their young much more than the fearful mothers. Critically, the docile mother's behavior not only has a direct effect on the behavioral and neural development of the young but, most remarkably, on gene expression itself, which is then propagated through multiple subsequent generations! In other words, a developmental outcome is not the product of a little bit of nature and a little bit of nurture; rather, it is the *joint* product of the coaction of completely interdependent endogenous and exogenous factors with neither having priority over the other.

David J. Lewkowicz

See also Experience-Dependent Plasticity; Infant Perception; Infant Perception: Methods of Testing; Perceptual Development: Face Perception; Perceptual Development: Intermodal Perception; Perceptual Development: Speech Perception; Theoretical Approaches

Further Readings

Blumberg, M. S. (2005). *Basic instinct: The genesis of behavior.* New York: Thunder's Mouth Press.

Gottlieb, G. (1997). *Synthesizing nature-nurture: Prenatal roots of instinctive behavior.* Mahwah, NJ: Erlbaum.

Lewkowicz, D. J. (2002). Heterogeneity and heterochrony in the development of intersensory perception. *Cognitive Brain Research, 14,* 41–63.

Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy of Sciences USA, 103*(17), 6771–6774.

Oyama, S. (1985). *The ontogeny of information.* Cambridge, UK: Cambridge University Press.

Wallace, M. T. (2004). The development of multisensory integration. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 625–642). Cambridge: MIT Press.

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology. Critical Periods Re-examined* [Special Issue]. *Evidence from Human Sensory Development, 46*(3), 233–234.

NAVIGATION THROUGH SPATIAL LAYOUT

Navigation is a nearly universal behavior for humans. It can take us from our hometown to Machu Picchu, and from across the room to our bed in the dark. To *navigate* means to plan and execute travel *through a spatial layout*. It often entails keeping track of our position and orientation as we move, and the ability to figure out where we are if we become lost. The "normal" range of human navigation ability is broad—some people learn new environments quickly and never seem to get lost, whereas others become disoriented easily, even in familiar places. This entry summarizes what is known about the psychological and neural aspects of navigation.

Psychological Aspects

Navigating over large distances is a complex process, sometimes requiring specialized instruments, maps, charts, or training. By contrast, navigating over smaller distances or in familiar territory may be accomplished without any specialized equipment or training. However, the basic requirements for successful navigation remain the same at large and small scales. At a minimum, navigators must first become oriented—that is, they must determine their current position and orientation with respect to a map of the environmental layout. The map could be physical (printed on paper, or perhaps displayed on a computer monitor), or it could be internalized (the navigator's memory of the environment). Then, navigators must plan a convenient route from the current location to the destination. Finally, navigators must keep track of, or update, their changing position and orientation while traveling in order to execute the planned route.

Most human navigation research has focused on situations in which the destination is not directly visible. These situations naturally fall into two types: landmark-based navigation (or piloting) and path integration (or dead reckoning).

In *landmark-based navigation,* we can see the immediate environment, but not the destination. This might happen because the destination is far away or hidden behind other objects. In this situation, we visually determine the distance and direction to prominent objects or locations in the environment ("landmarks"); if we can see a landmark in the real world that is represented in our map, we can use the landmark to determine our position and orientation with respect to the map. For example, when driving to a new restaurant, we might navigate by following a simple sketch map containing landmarks such as "green house on the left" and "second stop sign." When we see the green house on the left while traveling, we can determine our current location and orientation on the map and use this information to guide our progress toward the destination.

In *path integration,* our body sends signals to the brain about how fast we are moving, and we use this to determine our position and orientation relative to our starting point. For example, if our muscles tell our brain that we are walking at 1 meter per second (m/s), and we have been walking for 5 s, integrating this rate over 5 s shows that we are now 5 m from the starting position. Although path integration can certainly be performed by consciously estimating one's velocity and using mathematical calculating to integrate velocity over time, our brain is also capable of performing path integration without our explicit awareness. When walking across the room in the dark, for example, we may not be explicitly aware of our speed and how long we have been walking, but nevertheless we continue to feel oriented with respect to the room, despite having changed our location. This sense of orientation is based upon path integration. In real-world navigation, path integration might be used along with landmark-based navigation as we move about with our eyes open. However, most research focusing on path integration excludes the possibility of landmark-based navigation by asking people to navigate to a nearby destination with their eyes and ears covered.

One drawback of navigating exclusively by path integration is that as we move away from a familiar location, any errors in perceiving our own speed tend to accumulate over time, and this means that our ability to estimate our position and orientation will become increasingly inaccurate and imprecise. These errors can be eliminated if new information about our location becomes available (for example, by touching or seeing a familiar landmark).

In humans, path integration has been studied in a variety of ways. The most common involves the blind walking (or visually directed walking) technique. In this method, a participant views a target, puts on a blindfold, and then attempts to walk to the remembered target location. Typically, pace counting is discouraged and sound cues are minimized. After seeing a target in a well-lit environment, the average person can walk quite accurately to target locations that were initially 20 m away or more, even with eyes and ears covered. The average performance rapidly deteriorates, however, if the path involves a mixture of straight segments and body rotations. In the triangle completion task, for example, participants are guided without vision along two legs of a triangle, and are then required to return to the origin by completing the triangle. Performance is typically rather poor in this task. By contrast, some animal species, such as the desert ant, are astonishingly good at navigating by path integration. These creatures have been shown to use path integration when foraging for food along complex paths hundreds of meters long.

Regardless of whether we use landmarks or path integration to determine our position and orientation, we must somehow coordinate our current viewing perspective with the orientation of the map in order to plan and execute travel in the environment. If we're facing west and we're looking at a map with north at the top, how can we align our viewpoint with that of the map? In principle, we could imagine facing north, or imagine holding the map so that west is straight ahead, but people often find these options difficult and make navigational mistakes when required to use such strategies. Physically rotating one's body or the map greatly facilitates the task of aligning the two reference frames. "You are here" maps in shopping malls, which are often misaligned with the perspective of a person standing in front of the map, provide a good example of this. It is not unusual to see people tilting their heads or rotating their bodies in an effort to align their perspective with that of the map.

Neural Aspects

Navigation draws upon elementary sensory signals (from our eyes, ears, and muscles), as well as more abstract cognitive processes (such as attention, learning, memory, motor control, and planning). The brain areas responsible for performing these tasks are distributed throughout the brain. However, several brain regions appear to be particularly important for navigation.

Evidence in both animal and human studies suggests that the posterior parietal cortex plays a crucial role in keeping track of body motions within a frame of reference aligned with the body. Ultimately, to allow planning and control of navigation through the environment, this body-referenced information must be integrated into a mental map centered on the environment. Evidence suggests that the retrosplenial region plays an important role in performing this coordination between reference frames. Structures in and around the medial temporal lobe (which receives inputs from the retrosplenial region) play several roles that are crucial for navigation, such as storing spatial information about the environment, determining one's current position and orientation, and keeping track of one's movements relative to the environment. An early source of insight into the role of the medial temporal lobe came from work by John O'Keefe and Lynn Nadel. They measured the electrical activity of individual neurons in the rat hippocampus, a region of the medial temporal lobe, as the animals explored a tabletop-sized enclosure. Populations of hippocampal cells tend to become active whenever the animals move into certain regions of the environment. The neuronal activity is associated with the rat's location in the environment, regardless of the path taken to get there. These so-called *place cells* respond when visual, auditory, and olfactory cues are eliminated, suggesting that the cells are sensitive to the animal's body motions; they also respond to visual features of the environment. For example, if a cell becomes active when an animal is near a green bottle on the north side of the enclosure, moving the bottle to the east side will result in the cell now becoming active when the animal moves into the east side, near the green bottle. Other cells in the rat, called *head direction cells,* become active whenever the animal's head points in certain directions relative to the environment. Together, place cells and head direction cells provide a means of encoding an animal's location and orientation relative to the environment.

Edvard Moser, May Britt Moser, and their colleagues recently have discovered cells in the medial entorhinal cortex (MEC) of rats, a primary input to the hippocampus, that appear to provide a neural basis for the rat's ability to keep track of its changing location using path integration. Unlike place cells, which become active whenever the animal moves into a single region in space, cells in the MEC become active when the animal enters many individual locations in the environment. The locations that trigger activity in each of these cells are arranged in a hexagonal gridlike pattern. The regularity of the grid pattern suggests that these *grid cells* are not responding to particular features or objects in the environment as the animal moves. Instead, they respond as if they are keeping track of the rat's movements according to its own internal "meterstick." Grid cells and head direction cells have been found in the MEC, as well as neurons that show both grid and head direction firing properties. All of these cells are influenced by the animal's running speed. Thus, the MEC is extremely well-suited to integrate direction and position information as the animal moves about and to update this self-motion information. This information is then passed to the hippocampus and presumably forms an input to the place cells.

Several lines of evidence suggest that a similar organization holds true in humans. For example, if the human brain structures engaged by navigation are organized in a similar way as in the rat, one would expect people with injuries in the posterior parietal cortex or medial temporal lobe to exhibit navigational deficits. There is abundant evidence that this prediction holds true. One would also expect these regions to become activated in neurologically intact humans as they engage in navigation, and again there is evidence that this holds true. For example, using positron emission tomography, Eleanor Maguire and colleagues compared activation when individuals actively navigated through a familiar virtual environment versus when they performed an arrow-following task in the same environment that did not require keeping track of their position and orientation. The researchers found that during active navigation, there was increased activation in the posterior parietal cortex and hippocampus, particularly in the right hemisphere.

John Philbeck

See also Attention: Spatial; Self-Motion Perception; Spatial Layout Perception: Neural; Spatial Layout Perception: Psychophysical; Spatial Memory; Visually Guided Actions

Further Readings

Aguirre, G. K., & D'Esposito, M. (1999). Topographical disorientation: A synthesis and taxonomy. *Brain, 122,* 1613–1628.

Golledge, R. G. (1999). Human wayfinding and cognitive maps. In R. G. Golledge (Ed.), *Wayfinding behavior: Cognitive mapping and other spatial processes.* Baltimore: Johns Hopkins University Press.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation by path integration. In R. G. Golledge (Ed.), *Wayfinding behavior: Cognitive mapping and other spatial processes.* Baltimore: Johns Hopkins University Press.

Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S. J., Frith, C. D., & O'Keefe, J. (1998). Knowing where and getting there: A human navigation network. *Science, 280,* 921–924.

McNaughton, B. L., Battaglia, F., Jensen, O., Moser, E. I., & Moser, M.-B. (2006). Path integration and the neural basis of the "cognitive map." *Nature Reviews: Neuroscience, 7,* 663–678.

Taube, J. S. (2007). The head direction signal: Origins and sensory-motor integration. *Annual Review of Neuroscience, 30,* 181–207.

NEURAL CODING

See Neural Representation/Coding

NEURAL PROSTHETIC SYSTEMS

The brain receives input from all the senses and can transmit signals to the spinal cord for the execution of body movements. However, injury or disease can deprive the brain of sensory input or can interrupt output pathways, resulting in motor paralysis. The field of neural prosthetics aims to restore motor or sensory function through output and input pathways other than the brain's normal channels.

Brain stem strokes and diseases, such as amyotrophic lateral sclerosis (Lou Gehrig's disease), can lead to the most severe form of paralysis, as patients can lose all movement ability and become "locked in." Cervical spinal cord injuries (at the neck) also disrupt communication channels between the brain and spinal cord and can lead to quadriplegia (paralysis of all four limbs). All these patients have functioning brains and can still think about making movements and form the associated motor intentions, but they cannot execute these intentions. Motor prosthetic systems (MPSs) aim to harness these intentions and thoughts by recording their neural representation directly from the brain and relaying them to algorithms trained to control external devices, such as robotic arms or computer software. Ideally, in patients with spinal cord injuries, these devices may restore movements of the patients' own limbs by reconnecting the brain to the body below the injury, acting as a bridge that bypasses the severed section of the spinal cord.

An MPS is a framework that encompasses devices or programs to (a) record neural signals, (b) interpret the neural signals, and (c) control external devices. Thus, this field is highly interdisciplinary, requiring research across many branches of medicine and engineering, including neurosurgery, neuroscience, computer science, robotics, materials science, and microfabrication. This entry focuses on signals recorded from inside the brain from areas related to movement planning and execution. Three areas in particular have been shown to contain useful signals for MPS: the motor cortex (M1), the parietal cortex, and the dorsal premotor cortex (dPMC). These regions form a network for visually guided reaching by transforming visual input about the location of the target into commands that move the limb toward the target. Studies have shown that this network is affected by the reward associated with a target and can participate in decision making. The viability of this network for prosthetic control is enhanced when one considers that these areas are also engaged by cognitive processes, such as movement imagination. This entry first describes the neural signals in these areas and discusses some of the studies that utilized these signals for prosthetic control. It also describes the clinical trials currently under way. Then, this entry considers some of the challenges that have to be overcome and incorporated into the future development of flexible and effective MPSs.

Motor Neural Prosthetics Systems Using Variables Decoded From the Motor Cortex

The motor cortex (M1) can actuate movement in any body part by transmitting signals to the appropriate spinal cord level. These descending pathways form the communication channels that convey volitional movement instructions to the body. For example, reaching toward an object activates populations of neurons in M1 that encode the trajectory of the hand in space. Normally, these signals are sent to the spinal cord and are executed resulting in hand movements along the encoded trajectory. Researchers have used this finding as proof that M1 contains viable signals to drive MPSs because populations of neurons that control the trajectory of the hand could, in theory, directly control the trajectory of a robotic arm or a computer mouse. Much progress has been made to read out movements or movement plans from the motor cortex of monkeys, and more recently humans. A multielectrode array is typically implanted in the arm region of M1 in monkeys trained to perform reaches or move a cursor on a screen with a joystick. The activity of neurons is then recorded while the monkeys perform the tasks, and a mathematical model is built that relates the activity of the neurons to the cursor movement (or arm movement). Thus, every time the population of neurons fire, the model attempts to improve its ability to predict the movement direction of the cursor (or arm) by learning the relationship between neural activity and the resultant movement direction. After a satisfactory model is obtained, the model output is connected directly to the computer controlling the cursor. At this time, the joystick is disabled, forcing the monkeys to control the cursor using their neural activity and the model. The model, in effect, becomes a proxy for the arm and joystick. Immediately after the actual joystick is disabled, the monkeys generate muscular activity in their arm as they "will" the cursor to its target. This activity quickly subsides, however, as the monkeys learn to successfully move the cursor using only their thoughts. The signal pathway is thus transformed from the brain to the spinal cord to the arm to the joystick to the cursor, then to the brain to mathematical model to cursor. Thus, communication channels can be reopened by allowing paralyzed patients to control a joystick (or a computer mouse) using their thoughts. Clinical trials in humans have shown that this is indeed the case.

Instead of joysticks, other researchers have trained monkeys to use their thoughts to move a ball in a three-dimensional (3-D) environment. The monkeys in these studies were able to use M1 activity to operate and learn in the virtual environment without eliciting any overt movements. A recent extension of this work demonstrated that a similar strategy can be used to control a realistic robotic arm in a self-feeding task. These experiments successfully showed that trajectories can be decoded in real time from M1 and that neural prostheses work was feasible.

Cognitive Neural Prosthetics Systems Using Variables Decoded From the Parietal and Premotor Cortices

The trajectory information in M1 guides the hand (or the cursor) toward a specific goal, the target of the reach. The goal information must be known before and during reach execution to ensure accurate target interception. For example, a cursor moving on a computer screen must land on the appropriate icon to activate the desired program, making the location of the icon the motor goal. Several studies have shown that the goal information is encoded by neurons in the absence of movements or during imagined reaches to real targets. The goal information has been labeled *cognitive,* as this encoding specifies what the movement will accomplish rather than the muscular activation of how to accomplish it. Decoding the goal signal suggests that we can choose the icon and activate it without guiding a computer mouse toward it. In other words: think and click. The goal and trajectory information can be used independently to operate distinct applications, or synergistically to increase the accuracy of decoded reaches.

Goal information recorded from the parietal and premotor cortices has been used in prosthetic experiments. Monkeys with implanted electrodes in these areas were trained to place a cursor at instructed locations on a computer screen using their thoughts. Typically, a small circular light indicating the target location is flashed on a screen for 300 milliseconds (ms). Monkeys were trained to think about reaching to this target (thus forming a goal) but were not allowed to execute the reach. Neural activity encoding the goal of the intended reach was readily available less than 200 ms after the appearance of the flash. A mathematical model previously trained to associate neural activity with target location decoded the neural signals to determine the goal. When the decoded goal corresponded to the instructed location, a cursor was illuminated at the goal and the monkeys were rewarded. Thus, the monkeys had to activate a light at the target by thinking about moving to that target. Unlike trajectories, where the cursor path is continuously specified, this method is discrete and generates a single prediction about the location of the goal. Monkeys successfully placed cursors at the correct screen positions using their thoughts, proving that cognitive signals from the parietal and premotor cortices, like motor signals from M1, are viable signals for neural prosthetic applications. Motor and cognitive signals complement each other, and their combination will likely prove necessary to model and restore complex behaviors.

Cognitive signals in the parietal cortex have also been shown to be modulated by the expectancy of reward. The ability of the models to predict the goal information increased when an animal expected a large reward volume or a preferred reward (orange juice instead of water). This is useful for prosthetic applications, as the change in activity may be used to selectively modulate the amplitude of the population. For example, a model can be trained under a variety of reward conditions, increasing the number of discrete states (screen positions) the population can encode. The reward signal may also be related to motivation. Thus, it may be possible to predict the mood of a patient by monitoring these neurons.

Other Useful Signals for Prosthetic Control

All the studies previously described used signals recorded from electrodes implanted inside the brain. These invasive methods increase the risk to the patient but yield high information rates, as they are capable of recording the activity of single neurons. Implanted electrodes can also record electrical signals labeled as local field potentials (LFPs).

These signals arise from the summed electrical activity of a large number of synapses and neurons around the electrode and are a potentially rich source of information for prosthetic applications. Importantly, they are easier to record than action potentials and contain information related to behavior that can complement that decoded from single neurons. The single neuron activity and the LFPs are directly correlated with behavior and can likely be used to restore complex movements. Thus, to move a robotic arm, patients need to imagine moving their own arms.

A safer alternative to invasive methods is to record electroencephalograms (EEGs) from outside the brain by placing electrodes directly on the scalp or skull. EEGs arise from the sum of currents of large cortical areas beneath the scalp but are not related to any direct behavior or movement. Thus, patients must learn some arbitrary association between the EEG and the desired control signal resulting in long training times (sometimes lasting months) and poor information yields. Subjects have been able to use EEGs to move cursors on computer screens in two dimensions (2-D), but the utility of EEGs as control signals for more complicated behavior is doubtful.

Current Clinical Trials

Two companies have implanted electrodes in the motor cortex of paralyzed humans and have shown that the proof of principle experiments performed in monkeys can indeed be taken to the clinic. Neural Signals, Inc., founded by Philip Kennedy, has implanted locked-in patients using a glass cone filled with proteins that encourage neural growth. Neuronal processes grew into the cone and their activity was used by patients to move a cursor on a computer screen. Although multiple cone implantations have not yet been attempted due to the imposing electronics, this type of implant is promising, as it overcomes many of the problems plaguing implant technology. Even more intriguing is the recent claim that processes from many different neurons can grow into one electrode.

In 2004, Cyberkinetics Neurotechnology, founded by John Donoghue and colleagues, received FDA approval to implant the M1 region of paralyzed patients using an array of 100 silicon electrodes with platinum tips. Thus far, two patients with spinal cord injury, a patient with amyotrophic lateral sclerosis, and a patient with brain stem stroke have been implanted. M1 signals were readily recordable years after the injury, and immediate voluntary control of M1 neurons was possible. This was the first demonstration in humans that neurons in M1 can remain active even years after paralysis. In addition, LFPs similar to those identified in monkeys were also recorded. Patients were immediately (ignoring the short time needed to train the models) able to operate software and open and close a robotic hand. These clinical trials are encouraging and prove that neural prosthetic control of external devices is possible even in patients paralyzed for many years. Clinical trials to use cognitive signals for prosthetic control are just underway.

Outstanding Issues

Despite the results previously outlined, MPS research is still nascent. Next is a discussion of a few outstanding issues that need to be resolved if MPSs are to gain widespread clinical use. The biocompatibility of electrode arrays (or probes), the need to decode volition to ensure autonomous systems, and why incorporating feedback into an MPS to reduce errors may be necessary to restore complex movements are discussed.

Biocompatibility of Neural Probes

Perhaps the greatest challenge facing the use of electrode implants is the resultant tissue response to the neural injury elicited by the insertion of the implants into the brain. Many researchers have reported signal deterioration soon after brain implantation. For these devices to be useful, they should ideally remain functional indefinitely to enable self-contained and self-governing MPSs. Signal deterioration has been attributed to the general immune activation of the brain in response to the presence of a foreign device. Insertion of the electrode ruptures the vasculature and destroys all neurons in its path. Eventually, scar tissue engulfs the electrodes, isolating them from healthy neurons. Two strategies that are currently being investigated to resolve these issues include methods to minimize the immune response and methods to encourage neural growth around the electrodes. These problems must be overcome before MPSs become a common therapeutic procedure for paralyzed patients.

Deciding When to Act

The proof-of-principle research conducted in monkeys has predominantly used relatively straightforward behavioral tasks conducted in highly constrained and overtrained environments. Subjects do not generate any self-initiated movements driven by perception or through volition but are typically instructed to move. Algorithms are then synchronized with this prior information, eliminating the need to decode volition. This is further complicated by the realization that we can think about reaching but may sometimes choose not to execute those thoughts (especially if retribution is imminent). Neuroscience findings has further complicated these issues as research in perceptual decision making has shown that signals correlated with decisions are distributed throughout the brain. Experiments that use a motor response (such as an eye movement) to indicate a decision have shown that the firing rate of neurons increases as sensory information favoring a decision accumulates. Once a threshold is reached, the decision is elicited. Potentially, MPS algorithms may need to determine beforehand an individual neuron's threshold for movement commitment. More practically, there may be temporal features in neuronal activity (e.g., between LFP signals and/or spikes) within or between areas that indicate reliably when a commitment to move is made. Regardless, the brain's strategy to represent potential motor actions demonstrates the need for a reliable "Go" signal that reflects the decision or commitment to make a movement. Incorporating this signal into an MPS will yield more spontaneous motor control aspects that are required to perform more complex and voluntary tasks in a natural setting.

Strategies for Feedback

Other ways that MPSs may be able to operate as stand-alone devices is to provide some form of feedback about performance to the patient. Feedback can be thought of in both high-level terms about the outcome of the action, and in low-level terms detailing, for example, movement trajectories. High-level visual feedback is usually not disrupted in paralysis, as patients can perceive the outcome of their thoughts. Usually, the brain uses proprioceptive and visual feedback to optimize the accuracy and speed of a movement. Although visual feedback usually dominates, it is unlikely to provide the resolved sensory feedback needed for more complex movements, such as skilled grasping, which integrates information across both proprioceptive and cutaneous sources. Some patients suffering form deafferentation diseases cannot button their shirts or perform other skilled movements, although their ability to make gross reaches remains intact. Methods that compensate for these deficits are necessary if neural prosthetics are to restore natural movements.

It has been suggested that feedback could be provided via extracellular stimulation during the execution of a task. This could add the controlled external device to a patient's body image, improving perception of the device. This would result in MPSs that combine sensory and motor neural prosthetic technology. Extracellular stimulation has been delivered directly into the primary somatosensory cortex of humans and can induce gross sensory percepts related to tactile sensation. Other studies have shown that percepts induced by stimulations cannot be distinguished from tactile stimuli. However, it remains an open question whether anything resembling a natural and dynamic percept can be induced via an array of stimulating electrodes. Alternatively, extracellular stimulation could be used as a teaching signal to induce or consolidate learning, but this necessitates identifying right from wrong (without a technician) if the MPS is to operate in an unconstrained environment.

Sam Musallam and Brian Corneil

See also Cochlear Implants: Technology; Intentionality and Perception; Microstimulation; Prostheses: Visual; Reaching and Grasping

Further Readings

Andersen, R. A., Burdick, J. W., Musallam, S., Pesaran, B., & Cham, J. G. (2004). Cognitive neural prosthetics. *Trends in Cognitive Sciences, 8*(11), 486–493.

Carmena, J. M., Lebedev, M. A., Crist, R. E., O'Doherty, J. E., Santucci, D. M., Dimitrov, D. F., et al. (2003). Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS Biology, 1*(2), E42.

Chase, V. D. (2006). *Shattered nerves: How science is solving modern medicine's most perplexing problem.* Baltimore: Johns Hopkins University Press.

DiLorenzo, D. J., & Bronzino, J. D. (Eds.). (2008). *Neuroengineering.* Boca Raton, FL: CRC Press.

Hochberg, L. R., Serruya, M. D., Friehs, G. M., Mukand, J. A., Saleh, M., Caplan, A. H., et al. (2006). Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature, 442*(7099), 164–171.

Musallam, S., Corneil, B. D., Greger, B., Scherberger, H., & Andersen, R. A. (2004). Cognitive control signals for neural prosthetics. *Science, 305*(5681), 258–262.

Santhanam, G., Ryu, S. I., Yu, B. M., Afshar, A., & Shenoy, K. V. (2006). A high-performance brain-computer interface. *Nature, 442*(7099), 195–198.

Schwartz, A. B., Cui, X. T., Weber, D. J., & Moran, D. M. (2006). Brain-controlled interfaces: Movement restoration with neural prosthetics. *Neuron, 52*(1), 205–220.

Velliste, M., Perel, S., Spalding, M. C., Whitford, A. S., & Schwartz, A. B. (2008). Cortical control of a prosthetic arm for self-feeding. *Nature, 453*(7198), 1098–1101.

NEURAL RECORDING

Neural recording is a method of observing neural activity. It has provided neuroscience with a picture of how the nervous system works.

The collective activity of many brain areas, some containing almost 100,000 neurons per cubic millimeter of tissue, allows the nervous system to mediate our sensory perceptions, thoughts, decisions, and movements. These neurons communicate with one another via synaptic connections, of which there are upwards of one hundred million in the same cubic millimeter. With so many neurons, so tightly packed together, how can we possibly observe their activity?

A major problem in studying neural activity is that in most animals, the brain is enclosed in a hard bony skull. This has led to various noninvasive approaches to record the activity of neurons, such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI). Although a major advantage of these noninvasive techniques is that they can easily be used on human subjects, they can only measure the aggregate neural activity of a large volume of brain tissue. Another limitation is that while neural activity occurs on millisecond time scales, the metabolic signals measured by some of these techniques are extremely slow by comparison. Therefore, at the current level of technology, noninvasive techniques are unable to capture the activity of single neurons at millisecond time scales because of their low spatial and temporal resolution.

It is possible to achieve much higher resolutions if the activity of single neurons is directly observed through optical imaging or electrical recording using very tiny electrodes (microelectrodes). These are invasive methods that require opening the skull or other bony compartments and exposing the brain. Optical imaging records the activity of neurons by capturing the light emitted by certain molecules in response to electrical or chemical changes during neural activity. Although its current limitations include the restricted depth of accessible tissue and the inability to be easily used in freely moving animals, optical imaging has the potential to surpass microelectrode recording in its ability to provide high-resolution neural recordings. At present, however, recording neural activity using microelectrodes has the highest spatial and temporal resolution.

To understand the neural basis of perception, neural recordings have been performed on a diverse set of species, each having certain advantages and disadvantages: insects and lower animals have accessible nervous systems, whereas mammals have similar brain structures to humans. Although neural activity can be recorded in the peripheral nervous system, this entry focuses on the use of microelectrodes to record the activity of neurons in the central nervous system.

Neural Recording Using Microelectrodes

Electrical activity of neurons can be recorded by placing a microelectrode inside or in close vicinity to a single neuron or multiple neurons. Microelectrodes are very thin electrical conductors made either from metal, commonly tungsten, or a hollow glass micropipette filled with a conducting solution. This method is used on tissue slices (referred to as in vitro recordings) and on both anesthetized and alert animals (referred to as in vivo recordings). Because neural tissue does not contain pain receptors, microelectrode recordings are also performed on alert human patients undergoing neurosurgical procedures.

From the earliest neural recordings by the British electrophysiologist Lord Edgar Adrian, one of the recipients of the 1932 Nobel Prize in Physiology or Medicine for work on the function of neurons, the action potential (commonly referred to as a spike) has been considered the currency of neuronal activity. Action potentials are all-or-none changes in a neuron's electrical potential of approximately 100 millivolts that last about one millisecond. They are produced by ionic currents moving through special channels made of transmembrane protein molecules. The resultant action potential is propagated along axons, finally making a synaptic connection with the dendrite of another neuron. When an action potential arrives at a synapse, it produces a change in the electrical potential of the recipient neuron referred to as a synaptic potential, which is a few millivolts in amplitude and lasts tens to hundreds of milliseconds. If synaptic potentials are large enough to cross a threshold, action potentials are produced in the recipient neuron. Thus, there are two types of electrical signals that can be recorded by microelectrodes: fast, large action potentials, and slow, small synaptic potentials.

Intracellular Neural Recording

The change in a neuron's electrical potential can be recorded in a variety of ways, illustrated schematically in Figure 1(a). The most accurate measurements are obtained by placing an electrode inside the cell and measuring the voltage difference relative to a reference electrode placed in the extracellular fluid. If the tip size is small enough (below a micrometer), a glass or a metal microelectrode can be carefully advanced directly into the cell's cytoplasm (sharp-electrode recording). Slightly larger hollow glass electrodes can be used to access the cytoplasm by applying suction at the point of contact (patch-clamp technique).

Both of these intracellular recording techniques eventually degrade the internal environment of the neuron, and recordings are usually maintained for less than an hour. Intracellular recording is most easily performed on tissue slices in vitro, because there is no movement due to respiration or motion of the organism. However, in vivo intracellular recordings on awake animals have occasionally been performed despite being extremely difficult to maintain. By viewing individual neurons in a slice of tissue under the microscope, it is possible to study neural circuits using multiple intracellular recordings from different neurons.

Extracellular Neural Recording

Although intracellular techniques have provided neuroscience with the most accurate understanding of the electrical properties of single neurons, most of what is known about the link between neural activity and perception comes from in vivo extracellular recording of electrical signals by placing a metal or glass microelectrode next to but not inside the neuron. The tip of an intracellular electrode is very fragile and easily damaged. Extracellular electrodes, by comparison, are usually larger and can more readily withstand insertion into the neural tissue.

As the soma of an average neuron is about 10 to 20 micrometers in diameter, it is possible to advance an extracellular electrode with a tip size of a few micrometers until it is near a neuron. In this situation, the action potentials generated by that cell dominate the recorded signal (referred to as single-unit recording). Sometimes, the tip of a microelectrode might be located in the middle of a group of neurons, capturing the action potentials generated by all surrounding cells (referred to as multi-unit recording). Extracellularly recorded action potentials have a magnitude of about 100 microvolts, much smaller than those measured using intracellular electrodes. This can make separating an action potential from other electrical noise challenging. Also, unlike intracellular recordings, synaptic potentials cannot be easily observed with extracellular electrodes.

Filtering an extracellular signal below 300 hertz (Hz) removes the fast fluctuations of spikes, leaving the slower, more sustained currents due to synaptic potentials. This so-called local field potential (LFP) is believed to show the collective input to a brain area in contrast to spikes, which may better represent the output of a brain area. Analysis of

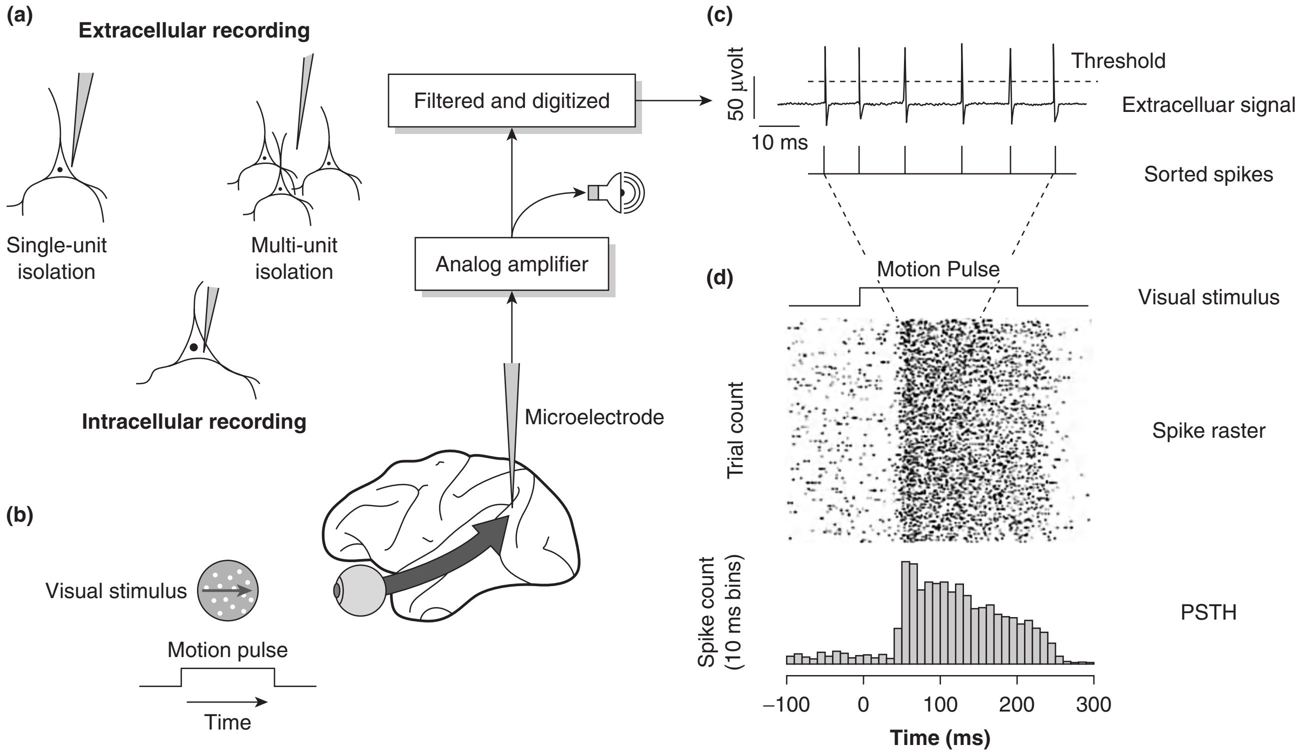


Figure 1 Recording Neural Activity With Microelectrodes and Spike Train Analysis

Notes: (a) Schematic of invasive electrophysiological recording techniques using a single electrode. (b) An extracellular in vivo recording experiment on a behaving monkey. A lateral view of the cortex is shown with the approximate electrode location. The electrode is not drawn to scale and is in fact smaller than a human hair at its tip. During the experiment and while neural signals were being recorded, the monkey viewed a visual motion stimulus presented on a computer monitor and performed a motion detection task. (c) The recorded waveform of the extracellular signal for a few spikes. From the extracellular waveform, the time of action potentials are determined by spike sorting. (d) The spike raster indicates spike times for each trial, aligned to the onset of the motion. The 200 millisecond (ms) motion pulse is indicated by the step. Temporal characteristics of the response of a neuron can easily be observed by averaging the response across trials in a peristimulus time histogram (PSTH).

LFPs in different frequency ranges has been used to address issues like perceptual binding, navigation, attention, time perception, and sleep.

Extracellular Neural Recording Example

In the example experiment shown in Figures 1(b) and (c), the activity of a neuron located in the visual cortex of a monkey performing a motion detection task was extracellularly recorded. The visual system has been a particularly strong focus of neuroscience because almost half of the primate cortex contains neurons that process visual information. Furthermore, visual stimuli are easily produced in a laboratory setting.

In this experiment, the animal subject was viewing a visual stimulus presented on a computer monitor. A single trial began when the monkey fixated a central point on the monitor and the stimulus appeared. The stimulus consisted of a patch of white dots moving on a gray background. The dots moved randomly at first, after which they all moved coherently in one direction for 200 milliseconds. They then moved randomly again for a short while before the stimulus disappeared. The monkey's task was to press a lever when the dots moved coherently to the right and the monkey was rewarded with juice for each correct trial.

While the monkey was viewing the stimulus and performing the task, a microelectrode was slowly advanced into an area of the visual cortex that processes motion stimuli. Because the raw signals from an extracellular electrode are very weak, they are first amplified several thousand-fold (Figure 1c). In many cases, the amplified electrode signal is played through a speaker so the experimenter can hear the neural signals. The timbre and loudness of the sound indicate the level of neural activity. For example, loud pops indicate a neuron is near the tip of the electrode. This qualitative auditory assessment as the electrode is advanced provides clues about the electrode's location. The electrode signals are then filtered and saved to a computer.

The recorded extracellular signal can be viewed as the sum of action potentials from the nearby neurons (within tens of micrometers), and background modulations resulting from synaptic potentials and action potentials of cells located farther away (within hundreds of micrometers). Spike-sorting is the process of extracting the time of each neuron's action potentials. This can be performed at the same time the data is being collected during the experiment, or later using the digitized waveforms that were recorded from the electrode.

Figure 1(c) shows an example of an extracellular recording that has a very high signal-to-noise ratio. In this case, it is easy to tell when spikes occurred based on the amplitude of the extracellular recording. Thus, any voltage modulation above a threshold value (shown by the dashed line) is considered a spike, and its time of occurrence is saved. Usually, however, spikes are not so clearly delineated from the noise. In such cases, more complex rules, like requiring the signal waveform to pass through one or more amplitude windows, are needed to select spikes. Sophisticated computer algorithms are also available for extracting spikes hidden inside noise.

The aim of these spike-sorting methods is to group the spike waveforms that have a similar shape and thus likely originate from a single neuron. If different waveforms are present, the recording is deemed multiunit (many neurons), whereas if all the counted spikes come from one dominant waveform, it is considered single-unit (single neuron). Therefore, a recording with a very high signal-to-noise ratio is probably a single-unit isolation, like the recorded neuron shown in Figure 1(c).

Properties of Spike Trains

Figure 1(d) shows a graph called a spike raster, where the time of each action potential is displayed by a tick mark, with each trial in a horizontal row. From the spike raster, it is easy to see that the number of spikes increases when the dots are moving on the visual display. Thus, the rate of action potential production, or the spike rate, of this neuron encodes visual motion. Also note that the increase in spike rate does not happen as soon as the motion occurs, but is delayed by about 50 milliseconds (latency), due to the time it takes for the neural activity to travel from the retina to the visual cortex. Notice also that the spike rate is initially very high, but quickly decays, although the motion stimulus remains constant. This is called *neural adaptation* and is a common property of many neurons that encode sensory information.

It can also be observed from the raster plot that the spikes produced by this neuron do not occur at regular intervals. A common feature of many neurons in the cortex is that the interspike intervals are highly variable. The raster also shows that there is variability from trial to trial in the total number of spikes produced in response to the stimulus (reliability). Furthermore, the spikes in different trials are not exactly aligned in time (precision).

To quantify the response profile of the neuron, it is customary to eliminate this trial-to-trial variability by averaging across trials. One way to do that is to count the number of spikes in consecutive time bins, for example every 10 milliseconds, and to plot the counts for each bin in a bar graph. With the resultant peristimulus time histogram (PSTH), it is easier to observe the average temporal characteristics of the neuron's spike response, such as the latency and adaptation previously mentioned (Figure 1d). It also shows more clearly that the cell had a small amount of baseline activity before the motion occurred.

Linking Neural Recording With Perception

Extracellular neural recordings can be readily made from animal subjects performing perceptual tasks. This experimental approach has provided neuroscience with the clearest understanding of the link between neural activity and perception. The example recordings in Figure 1 are from an alert monkey trained to respond to a visual motion stimulus. Although not shown, several behavioral parameters were simultaneously recorded with the neural recordings, including task performance, reaction time, and eye position using a video camera.

With both the behavioral and neural data in hand, a number of analysis techniques become available to answer several fundamental questions. For example, was the recorded neural activity used by the animal in forming its perception of the moving dots? If so, how was the spike activity of the recorded neuron processed by the rest of the animal's brain to create the perception of motion? How many neurons near the one recorded also contributed to the monkey's perception of the moving dots? And how do higher-level cognitive factors, such as attention, influence the link between neural activity and perception? Without the ability to record the neural activity, we would not be able to understand the neural basis of perception.

It is also possible to stimulate neurons and evoke activity by electrical or optical means. These microstimulation techniques complement neural recordings by providing another tool for linking neural activity to perception when performed on alert animals engaged in a behavioral task.

Multielectrode Recording Techniques

Because the brain relies on the response of many neurons to perceive the world (population coding), researchers have been interested in recording the activity of many cells simultaneously. Using two or more single electrodes to simultaneously record more than one neuron is possible, but technically challenging. For accessible cortical areas, planar and three-dimensional (3-D) multielectrode arrays containing 50 to 100 electrodes have been used to record from large neuronal populations. Linear array multielectrodes for simultaneously recording in different cortical layers have also been used, which can estimate the extracellular currents underlying field potentials using current source-density analysis.

A Practical Application

In recent experiments with paraplegic humans, a robotic arm was controlled by real-time decoding of recorded brain activity, either invasively by multielectrode single-unit recording or noninvasively by EEG. These studies show great promise for developing artificial limbs that can be controlled directly by the neural activity of the patient. Furthermore, by microstimulating a neural pathway with the correct spatial extent and temporal pattern using multiple electrodes, neural prosthetics (such as the cochlear implant) are paving the way toward artificial sensory interfaces predicted by popular science fiction.

Erik P. Cook and Navid G. Sadeghi

See also Brain Imaging; Evoked Potential: Vision; Magnetoencephalography; Microstimulation; Neural Representation/Coding; Prostheses: Visual

Further Readings

Dyan, P., & Abbott, L. F. (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems.* Cambridge: MIT Press.

Johnston, D., & Wu, S. M. (1994). *Foundations of cellular neurophysiology.* Cambridge: MIT Press.

Lemon, R. (1984). *Methods for neural recording in conscious animals.* Chichester, UK: Wiley.

Logothetis, N. K. (1999). Vision: A window on consciousness. *Scientific American, 11.99,* 69–75.

Nicolelis, M. A. L. (2008). *Methods for neural ensemble recordings.* Boca Raton, FL: CRC Press.

Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience, 21,* 227–277.

Salzman, C. D., Britten, K. H., & Newsome, W. T. (1990). Cortical microstimulation influences perceptual judgments of motion direction. *Nature, 346,* 174–177.

Velliste, M., Perel, S., Spalding, M. C., Whitford, A. S., & Schwartz, A. B. (2008). Cortical control of a prosthetic arm for self-feeding. *Nature, 453,* 1098–1101.

NEURAL REPRESENTATION/CODING

The activity of neurons ultimately relates to the world outside the nervous system. The perceptions, planned actions, and beliefs of an animal throughout its life are reflected in this neural code. The nature of this correspondence allows animals to act and react in ways that machines cannot come close to by being adaptable, robust, flexible, and fast. This correspondence between the outside world and inner neural code also provides a way of testing these neural responses by relating them to stimuli and actions in experiments. This entry reviews the triumphs and problems in interpreting this neural code, the wide array of experimental data, the technical and theoretical limitations, and ultimately how this information can be distilled. To ground the concepts covered with specific examples, this entry focuses primarily on one particular area of the mammalian brain, the primary visual cortex (V1), with some references to neural coding in the retina. Multiple levels of analysis are considered; for example, neural coding can be approached from a low level by noting how neural spikes relate to specific stimuli, like patterns of light in a visual image, or from a high level by understanding why one would want a neural code with those properties. The collection of analysis techniques provides insight to help bring the vast amount of experimental data into a more unified, coherent framework.

The Basics of Neural Coding

Neurons communicate throughout the brain using spikes of activity, known as *action potentials,* on the order of one millisecond in duration. These spikes are caused by a transient change in ion concentrations across the cell membrane. This leads to a change in membrane voltage that can be picked up by electrodes inserted near the neuron. Although there are methods of measuring neural response collectively (such as functional magnetic resonance imaging [fMRI], electroencephalography [EEG], optical imaging) single-unit electrophysiology offers an accurate means of measuring individual neural response both in time and space. For example, imagine trying to understand human speech by listening to a crowd of voices. EEG measures electrical impulses from the scalp. It is temporally precise but averages over space—like a microphone above a crowd. The analogy for fMRI, which measures changes in blood flow to a particular area following neural response, would be moving the microphone closer to a smaller group, but giving the average sound level every second. Single unit electrophysiology, in contrast to these other techniques, places a good microphone next to a particular speaker. It may not tell everything about the crowd (or even the other half of the conversation) but it gives detailed information about the individual person/neuron.

A full understanding of neural representation involves relating this response to the outside world. Early studies by Stephen Kuffler and others demonstrated that neurons in visual areas have receptive fields: regions of visual space where patterns of light can influence the firing of the neuron. The receptive fields of retinal ganglion neurons (whose axons make up the optic nerve) typically demonstrate a center-surround organization, where a central region might excite (or suppress) a neuron, while a surrounding region would do the opposite—these neurons respond best to differences in light levels.

Later work by David Hubel and Torsten Wiesel measured the receptive fields of V1 neurons, which have significantly different responses to stimuli. These neurons fire more strongly to particular orientations of lines, along with other stimulus features. One class of neurons (simple cells) was shown to have oriented regions of alternating lighting preference. In contrast to retinal ganglion neurons, which have a circularly symmetric, center-surround structure, these V1 cells find the difference between two or more nearby elongated regions. Later work mapped these receptive fields and showed how the particular pattern of light and dark preferences could be fit by a particular mathematical function, a two-dimensional (2-D) Gabor function—the details of which are not discussed here. A simple cell's selectivity for particular stimulus features, such as line orientation and position, are evident from a map of its subregions, or equivalently, the mathematical parameters of the Gabor function representation. However, it is well understood that such models are idealizations that account for only a portion of the neural response in these cells, and later visual stages are more difficult to characterize.

Characterizing Neural Responses

Using spots, bars, and gratings as stimuli is helpful, as they can be fully described by a small set of numbers—such as position, orientation, and size—but often these stimuli do not provide enough variation in examples to fully probe how stimuli can affect the behavior of the neuron. Unstructured noise stimuli (like the "snow" on an old TV set that wasn't tuned properly) can also be presented to a neuron and related to the resulting neural response. For example, if all the random stimuli that produced a neural spike are collected and then averaged together, the result is called the *spike-triggered average* (STA). This produces a receptive field as the simplest method of reverse correlation. However, such models require a great deal of data and have only a limited ability to characterize a neuron's response; because of this, it is clear that the response to a set of simple stimuli does not provide a straightforward prediction of how the neuron will respond to more complex stimuli.

There is a general tradeoff when characterizing the stimuli that affect a neuron's response. On one hand, structured, artificial stimuli (dots, bars, gratings) make implicit assumptions about neural response; they are often designed to elicit strong responses and require less data for analysis, but ultimately are less likely to generalize to new stimuli. Techniques that use random noise stimuli (reverse correlation) can make fewer assumptions by presenting a theoretically unlimited stimulus set, but often get weaker neural responses and can require more data than is possible to collect. Unfortunately, the tradeoff is ultimately a statistical one between restricting the stimuli or requiring massive amounts of data. One of the promising solutions to this tradeoff is the choice of natural stimuli. For example, in vision, the stimuli can be randomly sampled portions of natural images. This limits the size of the data set substantially (in comparison to noise stimuli) while probing the space more fully (in comparison to dots, bars, and gratings). Also, visual systems have evolved to respond to natural scenes; so many of the responses may be weak without the appropriate choice of stimuli. Samples from natural images are also not as well controlled or defined as artificial stimuli—making their use in many experimental paradigms less palatable. However, they do provide a distinct statistical advantage when a full characterization in necessary.

Efficient Coding

As previously noted, a number of methods are used in an attempt to characterize "what" causes a neuron to respond. Such methods can give us succinct mathematical descriptions that offer some predictive value for individual neurons. However, as the mathematical models become more complex, it becomes more difficult to understand the behavior of the neuron in a coherent way. Even if we could fully describe and predict the response behavior of a neuron, the question remains: "Why" does the neuron respond in that particular way? For our visual example, why do V1 simple cell receptive fields have the particular pattern of light and dark preferences (2-D Gabor functions)?

The ecological, efficient coding approach states that the goal of sensory processing is to efficiently represent the information that is behaviorally relevant to the animal. In the case of V1, the incoming visual information is coming from the natural world. There are particular properties of natural images that would suggest some codes are better than others. For example, light intensity often only changes at contours, so encoding primarily those changes by responding to lines/edges would be more efficient. What occurs in the left eye correlates with what occurs in the right eye, so receptive fields in each eye should be related for a particular neuron. In general, natural images are highly redundant, and removing these forms of redundancy in natural images would allow animals to use the information efficiently. Work by a variety of researchers, discussed next, has demonstrated that many goals of the early visual system directly relate the behavior of these neurons to the mathematical properties of natural images.

Our environment is highly predictable—neighboring pixels in natural scenes are highly correlated. One first step in efficient processing is to remove such redundant information. For example, Joseph Atick and colleagues proposed that individual ganglion cells remove these correlations so that the maximum amount of information can be passed through a constraining bottleneck—the optic nerve. However, when the visual signal reaches V1, there are more than 100 times as many neurons in primate V1 as axons in the optic nerve. Clearly the signal is not required to be compressed at this stage. What other type of efficient coding is useful at this stage?

One can place coding strategies along a spectrum from local, *grandmother cell* codes to distributed codes. A local code uses one neuron or relatively few neurons to represent a single, relevant piece of information. The traditional example is a grandmother cell code where the firing of one particular neuron represents information, such as whether or not your grandmother is present. Such neural responses would be easy to learn from and react to (approach or avoidance, for example), but clearly this code has disadvantages. For example, there are not enough neurons in the brain to represent every potential combination of visual features. On the other extreme, a distributed code uses many neurons to represent a single, relevant piece of information. For example, *compression* strategies can often result in highly distributed codes because part of the goal is to fully utilize the response range of every neuron. Taken to the extreme, a fully distributed code would be unreasonable in the brain as learning from and decoding such representations can be cumbersome. It would be difficult to respond to a neuron's firing if you need to sample input from every other neuron to interpret what that response means. Of course, depending on the input, a code can be local in some cases and distributed in others. For example, a pixel representation is local for representing small dots, but distributed for almost anything else. What was found to be the case in V1, and also in many other areas of the brain, is that the neural code predictably falls between these two extremes for natural stimuli.

David Field, Bruno Olshausen, and colleagues have shown that the principle properties of V1 simple cell responses can be explained by a strategy known as *sparse coding,* where only a few neurons fire strongly and the rest fire weakly or not at all. In one study, a neural network was used to search for the sparsest representation capable of representing natural images. Simulated neurons derived from such an approach mathematically resembled the receptive fields of real V1 neurons. Tony Bell and Terrence Sejnowski applied a similar efficient coding goal to natural scene data by increasing the amount of independence between neural filters using a technique called *independent components analysis* (ICA). Although a thorough explanation is beyond the scope of this entry, sparse coding and ICA actually use quite similar approaches, and their respective algorithms can be related mathematically. This is one case where multiple, high-level objectives may be working simultaneously with similar results, guiding the evolution and adaptation of the animal visual system.

The efficient coding approach argues that the ultimate goal of any neural representation is to be useful ecologically. The type of representation should increase the animal's evolutionary fitness. However, more explicit and immediate objectives, such as compression, sparse coding, or independent coding, can appear somewhat removed from this ultimate goal without explanation. Compression in the retina is beneficial given the limited size of the optic nerve. Such a strategy allows more data to be sent with fewer axons. The V1 coding objectives are efficient representations because they make the relevant information in the images explicit. That is, neurons are representing the relevant sparse or independent causes of the data, and with such explicit representations, it can make learning new associations simple. Other ecological reasons can exist for these objectives. It has also been argued that approaches like sparse coding are more metabolically efficient. With the correct constraints, a network of spiking neurons with a relatively small number of active neurons will use less energy than a network in which multiple neurons fire at intermediate rates. Explicit representation and metabolic efficiency are ecological justifications for the specific strategies mentioned. Whatever intermediate coding objective is considered, it must ultimately relate to ecologically relevant objectives.

Rate/Temporal Codes and Population Codes

Neural codes are more complicated than the simple individual neuron firing rate picture previously alluded to. For example, there is an ongoing debate about the extent to which the neural code is a rate code (only the average firing rate of a neuron matters) or a temporal code (the relative timing of the neural spikes is also important). These two strategies actually represent two endpoints on a continuum for analysis. If the window of time used for averaging is large, one is effectively assuming a rate code. However, if the time window decreases (say on the order of one millisecond or less), then only individual spikes are counted, giving a binary representation of a neural spike train. It is necessarily the case that more information is represented in finer windows, but interpreting the neural response becomes more difficult. The experimenter often must decide between the need for a great deal of data to analyze timing events and the pooling or binning of data needed to gain statistical power. What is clear is that many neurons are providing more information regarding the stimulus than that described by the firing rate.

In order to even establish a rate code, experimenters often average a neuron's response to a repeated stimulus, but, in many situations, the organism does not have the opportunity to perform such averaging. In many cases, responses of a population of neurons can be collected, and an immediate decision can be made based on those responses. However, as with temporal versus rate codes, there is often a tradeoff between making fewer assumptions and losing statistical power. A complete statistical treatment would consider the entire ensemble of neural responses in the decision. For example, a feature may be encoded by the relative timing of two neurons, requiring both neurons to be measured to describe the correct neural code. In practice, however, describing just one neuron's full response profile can be daunting. For this reason, population coding has proved most successful for only very well characterized areas, like the retina. As was the case for rate versus temporal codes, potential information is clearly discarded in simple averaging/pooling procedures. Neural coding is necessarily more complex than the rate-averaged single-unit level interpretation most often referenced in the literature. However, the question of how much information is carried by the relative timing of the population remains an active area of research.

Nonclassical Effects

In addition to the limitations on precision from temporal and population coding, individual neurons are also subject to a wide variety of nonclassical effects—effects that are not typically modeled by the standard methods. The number of deviations from the standard models can appear daunting and limitless without a unifying framework; the list includes the following: surround effects, contrast gain control, cross orientation inhibition, nonlinear spatial summation, nonfourier envelope responses, spatial frequency inhibition, supersaturation, and nonspecific suppression. Most of these effects are highly dynamic, inferential, and integrative, making them very difficult to model adequately. It is also difficult, or perhaps even misleading, to separate these effects. They can, however, be placed on a continuum between stimulus-driven and top-down effects. For more stimulus-driven effects, many fall outside the typical time frame and spatial extent of analysis. For example, stimuli normally elicit responses in V1 approximately 40 to 50 milliseconds (ms) after stimulus presentation, but light level, color, or contrast adaptation can have effects that last tens of seconds. Similarly, the area that can affect a responsive neuron is often much larger than the receptive field mapped out with simple stimuli. The number of nonclassical responses in both time and space within and outside the typical window of analysis is beyond the scope of any experiment and certainly beyond the scope of this entry. The search for nonclassical stimulus-driven effects is useful, and any well-formed general theory should eventually explain many of them. However, what is lacking in this area is a better theoretical explanation that makes sense of all these effects, rather than simply adding to the list with additional examples.

The effects that are more internally driven fall under many different interpretations—feedback, attention, or top-down influence. The neural code can adjust based on the goals or collective experiences of the animal. Attention directed at specific features or locations can affect neural responses in areas as early as V1. This modulation comes from feedback in higher brain areas, including areas associated with other sensory modalities. Cues to direct attention are often implicit, but can be explicitly regulated in practice using arrows, auditory instruction, or other learned contingencies. In this way, V1 is not a purely visual area, considering that the precise neural code can be affected by nonvisual cues. To process the vast amount of information efficiently, the neural code must adapt to cues outside the classical temporal and spatial extent, and top-down influence provides another ability to better encode what is most important to the animal.

The Future of Neural Coding

Although the examples used in this entry involve the early visual areas, many of the universal techniques to interpret the neural code were discussed.

Some of the tradeoffs will always be present—for example, the tradeoff between increasing statistical power with more specific stimulus assumptions (using structured artificial stimuli, such as dots, bars, gratings) and losing statistical power with increased stimulus generality (noise stimuli in reverse correlation) occurs in every sensory modality and method. Characterizations using natural scenes as stimuli appear to strike an appropriate balance by providing a wide array of ecologically relevant data that efficiently probes the space of neural responses relevant to the animal. Even simple cells in V1 are quite complex and require mathematical descriptions that are more involved than traditional models, and better characterizations are necessary for a more thorough understanding.

Computational objectives, such as compression, sparse, and independent coding, offer succinct, high-level explanations of why these neurons have particular responses to stimuli. Of course, neural response characterizations and computational explanations are often idealizations. The neural code is not simply a set of isolated neurons firing at particular rates, but rather a complex combination of neurons firing at precise times in a population. Although many new methods of characterization may be quite sophisticated, they are strictly limited by the amount of data available; this limits model complexity, temporal and spatial extent of receptive fields, as well as temporal and population resolution. Also, attention and feedback modulate the neural code to selectively process the massive amount of information previously mentioned. The effects of attention make any narrow interpretation of sensory coding give way to a high-level, multimodal, and at times psychological understanding.

Making sense of the neural code is not a simple process. A wide variety of new methods and models have been introduced in recent years to address the apparent complexity of neural activity. The introduction of ideas regarding ecology and efficiency provide the opportunity to reduce this apparent complexity. They introduce a unifying framework that integrates many disparate aspects of neural coding. Neural response properties are linked to ecological goals through the study of statistical properties of the environment. This new framework allows researchers to address both "what" neurons respond to, and more importantly "why" they respond that way. In this way, physiological, computational, and ecological results all become part of a more coherent picture of methods in neural representation and coding.

Mark V. Albert and David J. Field

See also Attention; Physiological; Auditory Processing: Central; Computational Approaches; Cortical Organization; Information Theory; Modularity; Multimodal Interactions: Neural Basis; Neural Recording; Psychophysical Approach; Reverse Correlation; Visual Processing: Extrastriate Cortex; Visual Processing: Primary Visual Cortex: Visual Scene Statistics

Further Readings

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J., Dan, Y., Olshausen, B. A., et al. (2005). Do we know what the early visual system does? *Journal of Neuroscience, 25,* 10577–10597.

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems.* Cambridge: MIT Press.

Doya, K., Ishii, S., Pouget, A., & Rao, R. P. N. (2002). *Bayesian brain: Probabilistic approaches to neural coding.* Cambridge: MIT Press.

Field, D. J. (1994). What is the goal of sensory coding? *Neural Computation, 6*(4), 559–601.

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience, 24,* 1193–1216.

NEUROPSYCHOLOGY OF PERCEPTION

The world surrounds us with an immense variety of sounds, colors, textures, and odors. In everyday life we need to receive and interpret this information to help us plan and adopt appropriate behaviors. Perception is the first phase of this interaction with the world. It progresses from initial encoding of data in a number of specialized sensory end organs for the different modalities (vision, hearing, taste, touch, and smell) through various stages of processing in subcortical and cortical structures in the brain, stages that elaborate, focus, interpret, and ultimately integrate this information to provide us with what we regard as our experience of the world.

Given the complexities of perceptual processing and the wide range of neural structures involved, it is not surprising to find that a large number of human diseases frequently impair perception. How these disorders—particularly those involving the brain—impact perception and what this can tell us about how perception works is the research domain of perceptual neuropsychology.

Neuropsychology is the scientific discipline that investigates the structure and function of the brain by assessing the effect of brain lesions in relationship to specific psychological processes and behaviors. This entry describes the basic procedures and goals of neuropsychology, neuropsychological studies of visual perception, and neuropsychology of other perceptual modalities.

The neuropsychological approach consists of assessing cognitive functions (i.e., memory, attention, perception, language, reasoning, etc.) by relying on tests specifically designed to assess each of those cognitive functions selectively. Neuropsychological tests also provide an accurate assessment of a selective cognitive function according to different materials and sensory modalities. In the context of perception, for instance, a neuropsychological assessment involves the evaluation of perceptual skills in visual, auditory, or tactile modalities, confined to different types of information such as faces, objects, sounds, colors, or more complex stimuli. At each test, the patient's performance is compared to the performance of a group of healthy individuals matched for age and education, providing an accurate estimate of the degree to which a selective function is impaired in a brain-damaged patient.

For many decades, the neuropsychological approach was the only method available to study where functions were localized in the human brain, and such studies were often performed in tandem with studies using more precise recordings or artificial lesions in animals, particularly monkeys. In the modern era, functional neuroimaging techniques have added powerful new tools to the study of the relation between structure and function in the human brain. However, such approaches also have their limitations, and the neuropsychological approach still provides us with information that cannot be obtained by other methods.

The two main approaches used in neuropsychology are the case study and the group study. Both have their advantages and disadvantages. *Single-case studies* with well-controlled experiments can provide important information about cognitive operations, but conclusions about the anatomic bases of the deficits in single cases is difficult. Natural lesions tend to be large, affecting a number of brain regions, limiting deductions about which region is the critical one. *Group studies* can show which lesions are associated with a specific functional deficit and which are not. On the other hand, group studies can suffer from the fact that any collection of patients is heterogeneous in many respects, regarding for example age, premorbid functioning, and lesion size and location. Careful selection, matching, and exclusion criteria can help, but heterogeneity can never be eliminated totally. Also, the assumption that all patients with a certain syndrome have the same functional deficit is not always correct, and group studies may inadvertently blur this distinction, leading to false conclusions about the anatomic correlates of function. To address some of the limitations previously described, neuropsychologists make use of an experimental technique named *dissociation,* which helps to identify the neural substrates of a selective cognitive function more reliably. In a *single dissociation,* the neuropsychologist is able to demonstrate that a specific brain lesion in region A is affecting the cognitive function X but not the cognitive function Y, which demonstrates that the two functions are independent and rely on different brain regions (a typical single dissociation is the inability to name an object when the patient can only see it, whereas he or she is able to name the same object by touching it). A *single dissociation* is strengthened when the neuropsychologist is able to demonstrate that a brain lesion in region A is affecting the cognitive function X but not Y and that a brain lesion in region B is impairing the cognitive function Y but not X. This is called a *double dissociation* and allows researchers to make conclusions about the independency of different neural substrates related to different cognitive functions (a typical example is a patient with a selective lesion in region A showing an impairment in visual memory but not auditory memory, and a patient with a selective lesion in region B showing an impairment in auditory memory but not in visual memory).

Goals of Neuropsychology

There are several reasons for using a neuropsychological approach to study any cognitive process. One of the earliest goals of neuropsychology was to correlate deficits in the mental processes of patients with the anatomic location of their lesions, the method of *anatomic-clinical correlation*. Using this logic, experimenters hope to gain insights into the neural structures that perform specific cognitive operations. If a lesion of a specific brain region results in a deficit in a certain behavior or percept, then the logical inference is that an important function that is required for that behavior or percept is performed by that cerebral region.

A second goal that has gained increased profile with the emergence of cognitive psychology as a field is to use psychological deficits as a *probe of the functional architecture of the mind* of how (rather than where) our brains process information. Many types of cognitive models have been developed to explicitly describe theories and mechanisms about the way our minds handle information. Common to many of these is the fractionation of a larger function into smaller modules that perform more limited and specific processes in a series of steps. In this context, a module is defined as a single component that is critical for the appropriate functioning of a more general cognitive skill. An example is the breakdown of face recognition into basic level encoding, perception of expression, perception of the facial elements that reveal identity, the matching of this identity information to memory stores of faces, and the accessing of semantic information about the person's name, biographical data, and so on. Which modules exist, the degree to which they exist independent of each other in the brain, and how information flows interactively between them are issues and proposals that can be constructed into a cognitive model, a functional architecture of how certain information is processed. Neuropsychological data can be a powerful means of testing whether these concepts are correct: If a patient's lesion selectively impairs a particular function and spares all other functions, this provides strong evidence for the existence of a module for that function. This linkage of deficits to functional architecture has similarities to the linkage of deficits to structural damage in the anatomo-clinical correlative method, but its aims are different and do not depend upon the presence of a focal or visible brain lesion.

Probing the structural correlates and the functional architectural implications of deficits are probably the chief scientific aims of neuropsychology. However, as befitting a discipline that sits at the interface between medicine and psychology, there are also goals of the discipline that are focused more on the disease process than on the psychological theories. Thus, a third goal is to *understand how various neurological or psychiatric disorders impact the cognition of the patient* suffering from the problem. This may be done by unraveling some of the basic mechanisms of diseases that are currently poorly understood. For example, studies showing that face perception is abnormal in at least some autistic patients have led some researchers to hypothesize that, given the importance of faces in our interactions with other people, the social dysfunction in this disorder may stem in part from derangements in perceptual mechanisms. On a more pragmatic level, this can also be done with the purpose of contributing to the diagnosis of a disorder or evaluating the risk of its development in the future. The patterns of memory, attentional, linguistic, and visuospatial impairments on a well-constructed neuropsychological battery may help clarify which of several types of dementia is present in an early stage in an elderly person with recent behavioral changes, or which patients with minimal cognitive impairment are at greater risk of going on to develop Alzheimer's disease over the next five years, for example.

A fourth and also clinically oriented goal that is gaining in interest is the potential role of neuropsychology in the *treatment and rehabilitation of neurological disorders*. Whether behavioral training paradigms have the ability to change the course of a patient's deficit remains a source of conjecture for many psychological deficits. At present, there are many such attempts being undertaken in the field of hemineglect (failure to process stimuli on the left side of space), and the impact of face-learning protocols are being evaluated experimentally in the treatment of prosopagnosia (the inability to recognize familiar faces) and autism. Even when the treatments being offered are not behavioral in origin, evaluations of their efficacy often are based in neuropsychological assessments.

The goals and the interpretation of data provided by neuropsychology are not without their own complexities and limitations. As with all disciplines, neuropsychology carries a set of assumptions about how the brain works. Whether these assumptions are valid for all cognitive operations can be questioned, and certainly all of them carry their own degrees of subtlety: few things are black and white. For instance, the ability to draw inferences about the functional and structural basis of human cognition from neuropsychological data requires that the mechanisms that the brain uses to process information do not radically change after brain damage. If a certain function X depends upon region A, then a lesion in A should cause a deficit M that reflects the loss of function X. However, if region B can learn to perform function X through some form of post-lesion plasticity (with either the same or a different type of computational mechanism than originally used by region A), then deficit M may not be apparent. In this case, the erroneous conclusion will be made that region A does not make an important contribution to the performance of X in the normal human brain.

Of course, there is evidence that the adult brain has some degree of plasticity and that things do change in the damaged brain. Cortical maps are reorganized after damage to peripheral structures, and even the opposite hemisphere appears to provide some compensation of loss of a lateralized function. These changes are somewhat limited, though, and it is generally assumed that these are minor quantitative variations rather than radically different qualitative rearrangements and shifts in computational mechanisms. Nevertheless, it is conceivable that for certain complex high-level perceptual processes, alternative routes and mechanisms to the same end might exist. In such a scenario, deficits may only emerge after a critical amount of damage has accumulated to disable not only the regions that perform the key computations in the normal brain, but also regions that might have been able to provide alternative compensatory routes or mechanisms for that function. The complexity that this violation of *pathological constancy* imparts to the interpretation of neuropsychological data is evident.

Neuropsychological Studies of Visual Perception

More is known about vision than any other perceptual modality: Hence, it is worth surveying what neuropsychological studies have contributed to our knowledge of visual function. The following is a short list of some of the more well-known neuropsychological syndromes of visual perception.

For over a hundred years, beginning with studies of soldiers in the Russo-Japanese War and World War I, it has been established that lesions of the occipital lobe lead to *cortical blindness,* with the location of the lesion corresponding to the area of the visual field lost (central versus peripheral, upper versus lower). Neurophysiological and neuroimaging studies have confirmed that the primary visual cortex (V1) is located in the banks of the calcarine sulcus, and that this serves as the primary cortical terminal of afferent visual information transmitted from the eyes through the optic radiations. From a neuropsychological perspective, visual loss from V1 damage conforms to many assumptions, including pathological constancy. The highly consistent relationship between the site of the lesion and the size and location of the blind area in the visual field on the opposite side of the lesion is well known and not an active area of research. More interesting is the question of what visual functions remain possible within these regions of blindness. Despite the denial of vision, some cortically blind patients can locate objects by pointing or eye movements, can sense the motion of objects, and can perform a number of other tasks at rates better than chance. This phenomenon is referred to as *blindsight,* and its investigation seeks to refine our notions of what visual operations are carried out independent of the processing in V1.

Beyond V1, an important anatomic concept that has grown from a convergence, first between neurophysiology and neuropsychology, and later from functional neuroimaging as well, is that of two visual processing streams, a dorsal one involving occipitoparietal cortical regions and a ventral one involving occipitotemporal regions, neural pathways connecting occipital and parietal regions and occipital and temporal regions respectively. These visual processing streams were initially conceptualized as a *dorsal "where" stream* for information about the spatial location of objects, and a *ventral "what" stream* for information about the identity of these objects. More recent alternative formulations suggest that the dorsal stream is involved in preparation for action and the ventral stream in object recognition. The debate about these formulations continues to be fueled by data from neuropsychological studies of patients with dorsal or ventral lesions.

Occipitotemporal lesions result in a variety of perceptual syndromes. Severe bilateral lesions may cause a *general visual agnosia,* in which the patient is unable to recognize objects (e.g., the man who mistook his wife for a hat). One of the aims of current neuropsychological research is to understand the types of functional disturbance that can lead to this severe deficit. Time-honored distinctions contrast an *apperceptive* with an *associative* form of agnosia. In the apperceptive form, although the patient can see, the perceptual information is so degraded that it is impossible to recognize what the object is. In the associative form, the perception information is accurate, but the knowledge about the perceptual properties of objects is impaired or inaccessible, rendering the perceptual input useless. Neuropsychological research has also established a possible third form, an *integrative* agnosia, in which perception of parts and features of objects is accurate, but the ability to integrate the parts into a coherent whole object is impaired, impeding recognition of the entire item.

Selective agnosias also exist. Some patients suffer from prosopagnosia. These patients do not have problems recognizing objects at a coarse level (they can tell faces from hats), but more sophisticated distinctions between different faces (and in some patients, between different hats or other objects) are very difficult for them. Again, research has established that there is not one type of prosopagnosia, but several variants, from damage to different anatomic structures of a face-processing network and to different functional components of a face-processing strategy. *Pure alexia,* in at least some patients, can also represent a selective agnosia, in which the ability to read words is lost, even though the patient has sufficient residual linguistic function to write, and can visually recognize other objects well. In other cases, pure alexia may represent a disconnection syndrome, due to damage transferring information from the visual cortex to linguistic processing centers in the left hemisphere.

Topographagnosia, the inability to find one's way in familiar surroundings, is another type of selective agnosia. It illustrates the complexities of functional modularity, because subjects navigating through space can use a variety of strategies. Different types of topographagnosia can result from damage to different orienting strategies. Some forms may be due to loss of the ability to recognize places and landmarks, and therefore a type of object recognition deficit: Not surprisingly, this type of topographagnosia is associated with damage to the occipitotemporal cortex. Other forms represent the inability to form a mental map of the environment, following hippocampal damage, or problems with utilizing a sequence of directions.

Achromatopsia, the loss of the ability to discriminate one color from another, is another deficit linked to occipitotemporal damage.

Occipitoparietal lesions do not impair object recognition, but result in a variety of syndromes involving spatial processing or the direction of attention into space. Though rare, defects in motion processing (*cerebral akinetopsia*) have been described. Bilateral occipitoparietal lesions are classically associated with *Bálint's syndrome,* which includes three separate features that are not always present in each patient. First is *simultanagnosia,* the inability to process more than one object in the environment, which likely represents a failure of mechanisms of attention. Second is *ocular motor apraxia,* in which the patient has difficulty initiating and targeting eye movements to spatial locations accurately. Third is *optic ataxia,* in which the patient's ability to reach for objects is impaired. Both ocular motor apraxia and optic ataxia represent difficulties in deriving spatial coordinates of items in the environment for the guidance of the subject's own motor responses.

Unilateral occipitoparietal lesions are usually less dramatic. Hemineglect is the classic deficit, usually from a right-sided lesion, in which a subject fails to notice stimuli on the left side of space. *Dressing apraxia* may represent an impairment of the ability to orient one's body to the complex and dynamic three-dimensional properties of one's own clothing.

Neuropsychology of Other Perceptual Modalities

After vision, the next most studied perceptual modality in neuropsychology is audition. Just as there is a primary visual area, there is also a primary auditory cortical area, in the temporal lobe, and damage to this region causes *cortical deafness.* This disorder is rare, presumably because auditory information from each ear is distributed to both hemispheres, and damage to both temporal lobes is a far more infrequent occurrence than bilateral occipital lesions that cause cortical blindness. There are also blindsight-like issues with cortical deafness, regarding what sorts of auditory abilities might still be present. Beyond primary auditory areas, damage to the auditory association cortex may lead to *auditory agnosia,* in which subjects can still perceive frequency, intensity, and duration of sounds, but cannot recognize sounds, speech, or music. More selective auditory agnosias include *word-deafness,* in which subjects are unable to comprehend speech, and *amusia,* in which subjects cannot recognize music, most likely through loss of the ability to process melody. Neuropsychological studies have gone further to suggest that there may be selective deficits in recognizing musical contour versus musical intervals in amusia.

In the somatosensory modality (touch), there are also deficits linked to damage to the primary sensory cortex in the precentral gyrus. As well, there are patients with *tactile agnosia,* who have intact perception of shape and texture, who cannot recognize objects by touch. This too has been fractionated into an apperceptive (or integrative) form and an associative form.

For the modalities of taste (gustatory system) and smell (olfactory system), there is much less information. It has long been known that damage to the orbitofrontal cortex or anterior and medial temporal lobe can impair the sense of smell, but further classification of central olfactory disorders has not yet been attempted. The primary gustatory cortex is located in the insula and adjacent inner operculum, which project to the orbitofrontal cortex, allowing for an integration of smell and taste. There is even less neuropsychological data on cerebral disorders of taste than there is for smell.

Giuseppe Iaria and Jason J S Barton

Acknowledgements: Giuseppe Iaria is supported by fellowships from the Michael Smith Foundation for Health Research and the Alzheimer Society of Canada. Jason J. S. Barton is supported by a Canada Research Chair and a Senior Scholarship from the Michael Smith Foundation for Health Research.

See also Agnosia: Visual; Attention: Disorders; Brain Imaging; Consciousness: Disorders; Vision: Developmental Disorders

Further Readings

Farah, M. J. (2004). *Visual agnosia* (2nd ed.). Cambridge: MIT Press.

Peterson, M. A., & Rhodes, G. (2003). *Perception of faces, objects, and scenes.* Oxford, UK: Oxford University Press.

Rizzo, M., & Ettlinger, P. J. (2004). *Principles and practice of behavioral neurology and neuropsychology.* Philadelphia: W. B. Saunders.

Vallar, G. (2000). The methodological foundations of human neuropsychology: Studies in brain-damaged patients. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology.* Amsterdam: Elsevier.

NONVERIDICAL PERCEPTION

A person's entire life experience—everyone, everything, every experience he or she has ever known—exists to that person only as a function of his or her brain's activity. As such, it does not necessarily reproduce the physical reality of the world with high fidelity. *Nonveridical perception* is the sensory or cognitive discrepancy between the subjective perception and the physical world. Of course, many experiences in daily life reflect the physical stimuli that fall into one's eyes, ears, skin, nose, and tongue. Otherwise, action or navigation in the physical world would be impossible. But the same neural machinery that interprets veridical sensory inputs is also responsible for one's dreams, imaginings, and failings of memory. Thus, the real and the illusory or misperceived have the same physical basis in a person's brain.

Types of Misperceptions

Misperceptions (that is, perceptions that do not match the physical or veridical world) can arise from both normal and pathological processes. Everyday perception in the normal brain includes numerous sensory, multisensory, and cognitive misperceptions and illusions. But these may also result from abnormal brain processes or physiological conditions, such as hypoxia, drug consumption, brain trauma, and neurological diseases, among others. This entry explores some types of misperceptions that occur in the healthy brain in standard physiological conditions.

Sensory Misperceptions

Sensory misperceptions are phenomena in which the subjective perception of a stimulus does not match the physical reality. Sensory misperceptions occur because neural circuits in the brain amplify, suppress, converge, and diverge sensory information in a fashion that ultimately leaves the observer with a subjective perception that is different from the reality. For example, lateral inhibitory circuits in the early visual system enhance the apparent contrast of edges and corners so that these visual features appear to be more salient than they truly are.

Visual Misperceptions

In a visual illusion, the observer may perceive a visual object or scene that is different from the veridical one. Alternatively, the observer may perceive an object that is not physically present, or fail to perceive an object that is extant in the world. In the scintillating grid illusion (a type of brightness illusion; Figure 1), the subject perceives an illusory darkening of veridical white circles at the intersections of a grid.

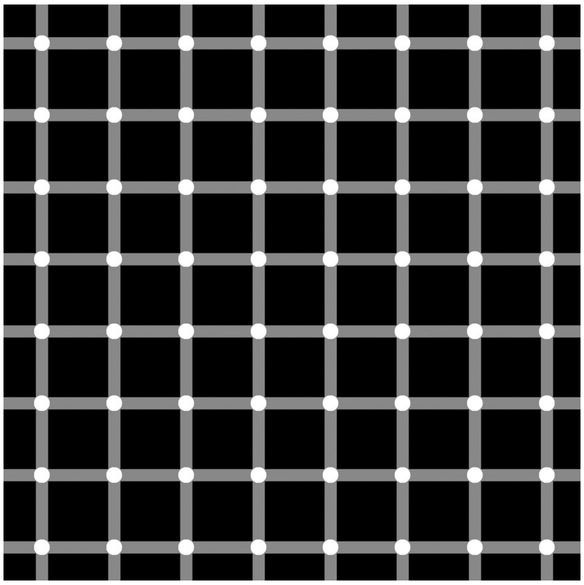


Figure 1 The Scintillating Grid Illusion

Source: Modified from Schrauf et al., 1997.

Notes: As you move your gaze around the image, observe the white circles on the intersections of the grey bars. Illusory dark dots seem to appear and disappear on top of the white circles, except for the one that you focus your gaze on at any given time.

Another well-known visual illusion is the perception of apparent movement. In this illusion, one object turns off while another object, in a separate spatial location, turns on. The perception is of a moving object that travels from the location of the first object to that of the second object. Movie marquees and motion pictures are practical applications of this principle.

Size illusions exemplify the importance of context in visual perception. In the Ebbinghaus illusion, a central circle will appear larger or smaller depending on the size of the circles surrounding it (the central circle will look large when the surrounding circles are small, and vice versa).

Errors in the perception of distance are another common type of visual misperception. Distances may be underestimated or overestimated, depending on various contextual cues. Dennis Proffitt and colleagues showed that people wearing backpacks consistently estimated distances as longer than people not wearing packs. More recently, Russell Jackson and Lawrence Cormack found that observers judged a cliff as higher when looking down from the top than when looking up from its base. Scientists have speculated that errors in distance perception such as these may have an adaptive value: For instance, a subject that overestimates the vertical drop of a steep cliff may descend it with great care, thus improving his or her chances of survival.

Auditory Misperceptions

In an auditory illusion, the listener may perceive sounds that are not present or that are different from those physically present. Auditory misperceptions are common while listening to speech embedded in noise (this is especially true for nonmeaningful speech; meaningful speech is easier to hear in noise). Other misperceptions involve speech segmentation. For example, consider the following two phrases: "How to recognize speech," and "How to wreck a nice beach." Although both sound patterns are almost identical, the context usually helps determine the correct perception.

The phantom words illusion, demonstrated by Diana Deutsch, is also related to speech perception. Here, the subject listens to recorded overlapping sequences of repeating words or syllables with different sounds presented in the left and right speakers. Even though the same exact sounds are repeated over and over, the listener will pick up specific phrases that appear to change from time to time. In one specific demonstration, the following phrases were heard by various speakers "no way," "when oh when," "mango," and "window." None of the perceived phrases were real, however, the listeners' brains simply made them up, as if trying to make sense out of meaningless speech.

Tactile and Pain Misperceptions

In tactile and pain illusions, the sensation of touch or pain differs in important ways from the physical stimulus. The thermal grill is an example of a powerful pain illusion, created by a grill of warm and cool interleaving bars. When a subject rests a hand against the grill, he or she experiences a sensation of burning pain. However, neither the warm nor the cool bars are painful to the touch when experienced in isolation. Thus, the thermal grill illusion further demonstrates that sensory perception does not work in absolute terms, but it is context dependent.

The cutaneous rabbit is a well-known tactile illusion. In the cutaneous rabbit, the subject's skin is sequentially tapped in two different locations. The subject's perception is that intervening (and nonstimulated) skin locations were also tapped, perhaps reflecting a kind of tactile "filling-in" process.

A number of tactile misperceptions are based on neural adaptation, that is, the decrease in responsiveness of sensory neurons when exposed to unchanging stimulation. To experience a tactile misallocation effect based on adaptation, close your eyes and try to touch the edge of one of your socks with your fingertips. It is very likely that you will fail to pinpoint the exact location of the sock's edge. The reason is that, in the absence of change, the somatosensory receptors signaling the edge of the sock get adapted (that is, they cease to respond). The perceptual consequence of such neural adaptation is a decrease in sensitivity. Adaptation effects are also common in other sensory modalities, such as vision, taste, and olfaction. For instance, adaptation helps explains why an unpleasant odor is more tolerable after a few minutes of exposure, or why the 10th bite of a delicious dish is never as good as the first one.

Gustatory and Olfactory Misperceptions

Except for those effects that are adaptation-based, taste and smell misperceptions appear to be less frequent than those of other sensory modalities. One powerful example of a taste illusion is the perceived localization of flavor (caused by olfactory input combining with taste) as occurring in the mouth. Tastants entering the mouth stimulate taste buds on the tongue, and vapors also reach the olfactory receptors through a passage that connects the mouth and nasal cavity. The stimulation of the olfactory receptors combined with stimulation of the taste buds create "flavor." However, we usually call this "taste" and associate it solely with stimulation of the tongue, presumably because of tactile sensations on the tongue. Another commonly experienced taste misallocation effect is the (incorrect) sensation of taste in areas of the mouth that have no taste receptors (for instance the roof of the mouth). This illusion may arise because the brain misinterprets the touch sensations of the food inside the mouth as the sensation of taste.

Multisensory Misperceptions

Multisensory misperceptions result from the interaction of two or more sensory modalities. The McGurk effect, or McGurk illusion, demonstrates that speech perception results from integrating both auditory and visual information. In this illusion, the visual stimulus is a video of a speaker making the lip and face movements for producing a sound such as /ga/. The auditory stimulus is a different sound, such as /ba/. However, a person watching the lip and face movements for /ga/ while receiving the auditory stimulus /ba/ actually hears a different sound, such as /da/.

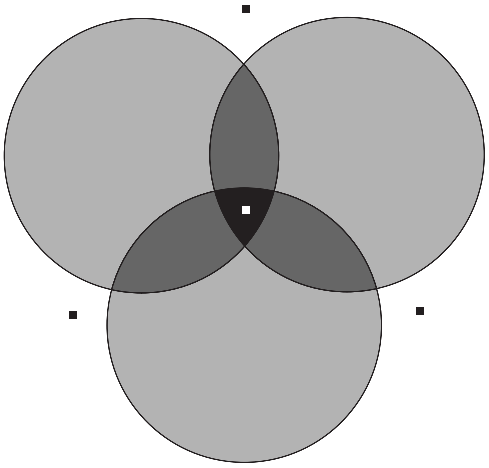


Figure 2 Perceived Brightness Is Affected by Voluntary Attention

Source: Modified from Tse, 2005.

Notes: Fixate your gaze on one of the fixation dots and direct your attention to one disk or another. The attended disk appears to darken in an illusory fashion, whereas the unattended ones appear unchanged.

Cognitive Misperceptions

Cognitive misperceptions involve higher-level cognitive functions, such as attention, memory, and causal inference.

Misperceptions Due to (In)attention

Attended objects may appear to be more salient or to have higher contrast than unattended objects (Figure 2); these perceptual effects have well-documented neural correlates in the visual system. Also, there are several cognitive neuroscience paradigms in which the allocation (or misallocation) of attention plays a critical role. These are described in the following sections.

Change Blindness. This is a phenomenon in which an observer fails to notice a large (and often dramatic) change within a visual scene. Such unnoticed changes usually take place during a transient interruption, such as a blink, a saccadic eye movement, or a flicker of the scene. Daniel Simons and Daniel Levin conducted a classical change blindness study, in which an experimenter approached a naïve pedestrian to ask for directions. While the pedestrian was providing directions, two additional experimenters carrying a door between them, rudely passed between the initial experimenter and the pedestrian. During this brief interruption, the original experimenter switched places with one of the door-carrying experimenters and continued the conversation with the pedestrian. Although the two experimenters looked different from each other and were dressed in different clothes, about half of the pedestrians failed to notice the switch. Change blindness is also common during changes that are introduced during cuts or pans in movies. Indeed, motion pictures often contain continuity errors (for instance, a glass of wine that is empty in one scene will be full in the next scene), despite very careful editing. Such continuity mistakes are usually unnoticed by spectators.

Inattentional Blindness. This is a phenomenon in which observers fail to notice an object that is fully visible in the display. Inattentional blindness differs from change blindness in that no memory comparison between the prechange state and the postchange state is needed: The missed object is fully visible at a single point in time. In a famous example of inattentional blindness, Daniel Simons and Christopher Chabris asked observers to watch a video showing people passing a basketball and to count how many times the members of a basketball team passed a ball to one another, while ignoring the passes made by members of a different team. While concentrating on the counting task, most observers failed to notice a person wearing a gorilla suit walking across the scene (the gorilla even stops briefly at the very center of the scene and beats its chest!). In this situation, no interruption or distraction was necessary, as the assigned task of counting passes was absorbing. Further, the observers had to keep their eyes on the scene at all times in order to accurately perform the task. Daniel Memmert showed, using eye-tracking recordings, that many observers did not notice the gorilla even when they were looking directly at it.

Memory Illusions

A large variety of memory illusions are commonly experienced by most people, such as déjà vu—the feeling that a novel situation has been previously experienced, or jamais vu (the opposite of déjà vu)—the illusion that a familiar situation has not been previously experienced. Another memory illusion is known in cognitive science as the *misinformation effect,* that is, the tendency for misleading information presented after the event to reduce one's memory accuracy for the original event. In a classic experiment, Elizabeth Loftus and John Palmer asked observers to estimate the speed of a car hitting another, after watching a video recording of a car accident. Observers that were asked how fast the car was going when it *hit* the other car gave lower speed estimates than observers that were asked how fast the car was going when it *smashed into* the other car. Experiments such as this demonstrate that memory illusions are critical to consider in the context of eyewitness testimonies. Eyewitnesses also generally report scenarios that are consistent with their expectations, a phenomenon known as *confirmation bias*. Such confirmation bias could also be responsible for some gross and tragic misperceptions, for instance in the police shootings of unarmed black subjects (the police officers declared to have "seen" a weapon, rather than a harmless object or an empty hand), possibly due to racial stereotypes associating blacks with violence and gun possession.

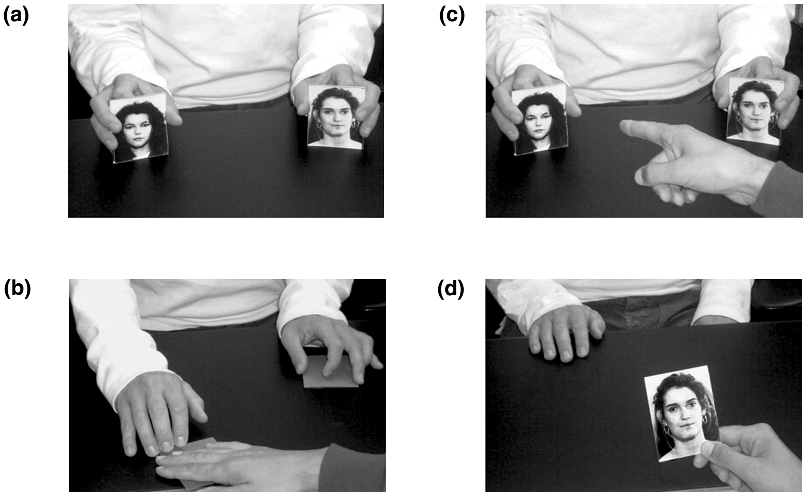


Figure 3 A Manipulated Trial During a Choice Blindness Experiment

Source: From Johansson, P., Hall, L., Sikström, S., & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science, 310,* 116–119. Reprinted with permission from AAAS.

Notes: (a) Subjects are presented with two female faces and asked to choose the most attractive one. Unknown to the subjects, a second card with the opposite face is hidden behind each of the visible faces. (b) The subjects indicate their choice by pointing at the preferred face. (c) The experimenter flips down the pictures and slides the hidden picture over to the participant, covering the original picture with the sleeve of his moving arm. (d) Participants pick up the picture and are immediately asked to explain why they chose the way they did.

Illusory Correlations

We infer cause and effect in everyday life. When A precedes B, we often conclude that A *causes* B. This causal inference is integral to our perception of magic tricks, many of which involve apparent violations of causality. A skilled magician will link two unrelated events, A and B, by making sure that event A (pouring water on a ball) always precedes event B (the ball disappearing). Although A does not *actually* cause B, the spectators will perceive the events as causally related. This type of illusion—seeing a correlation that is not there—is termed an illusory correlation.

Free Will and the Illusion of Choice

We live our lives under the practical assumption that we are free to make our own choices. However, the existence of free will is debatable. More often than not, we are unaware of the exogenous and endogenous constraints that explain our choices. Thus, our experience of free will may be no more than a sophisticated cognitive illusion. A recent experimental paradigm moreover demonstrates that we can be blind to the outcomes of our choices. In this paradigm, developed by Petter Johansson and colleagues, the relationship between a subject's choice and its outcome is surreptitiously manipulated. Subjects were shown picture pairs of female faces (Figure 3a) and asked to choose which face in each pair they found most attractive (Figure 3b). On some trials, participants were also asked to verbally describe the reasons behind their choice. Unknown to the subjects, the researchers occasionally switched one face for the other (Figure 3c), after the subjects made their choice. During manipulated trials, the picture revealed by the experimenter as the one that had been chosen became the opposite of the subject's initial choice (Figure 3d). Interestingly, only 26% of all manipulated trials were caught by the subjects. But even more surprisingly: When the subjects were asked to state the reasons behind their choice in the manipulated trials, they confabulated to justify the outcome, which was opposite to their actual choice. Johansson and colleagues called this phenomenon *choice blindness*. Choice blindness shows that there are certain situations in which we do not perceive the difference between what we get and what we originally asked for. The potential implications of this phenomenon for everyday life are intriguing, especially as follow-up experiments have started to show that choice blindness is not restricted to vision, but it extends to other sensory modalities.

Is Everything an Illusion?

In the *Matrix* movie, Morpheus explains to Neo that everything we consider "real"—that is, all that we perceive through the senses—is the product of neural activity in the brain. But what the movie does not say is that, even when Neo awakens from the illusory world of the Matrix into the veridical world, his brain will continue to construct his subjective experience—as all of our brains do—and this experience may or may not match reality, to varying degrees. Thus, in a way, we all live in the illusory "matrix" created by our brains. Years before the *Matrix* movie, neurologist and Nobel laureate Sir John Eccles wrote that the natural world contains no color, sound, textures, patterns, beauty, or scent. Thus, color, brightness, smell, and sound are not absolute terms, but subjective, relative experiences that are actively created by complicated brain circuits. This is true not only of sensory perceptions, but of any other experience. Whether we feel the sensation of "redness," the appearance of "squareness," or emotions such as love or hate, these are constructs that result from electrochemical impulses in our brain.

Susana Martinez-Conde
and Stephen L. Macknik

See also Aftereffects; Atmospheric Phenomena; Attention: Effect on Perception; Attention and Consciousness; Attractiveness; Auditory Illusions; Change Detection; Contrast Enhancement at Borders; Consciousness; Consciousness: Disorders; Cultural Effects on Visual Perception; Emotional Influences on Perception; Hallucinations and Altered Perceptions; Lateral Inhibition; Multimodal Interactions: Visual–Auditory; Out-of-Body Experience; Phantom Limb; Vision: Developmental Disorders; Visual Illusions

Further Readings

Johansson, P., Hall, L., Sikström, S., & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science, 310,* 116–119.

Levesque, R. R. (2006). *The psychology and law of criminal justice processes.* New York: Nova Science.

Macknik, S. L., King, M., Randi, J., Robbins, A., Teller, T. J., & Martinez-Conde, S. (2008). Attention and awareness in stage magic: Turning tricks into research. *Nature Reviews Neuroscience, 9*(11), 871–879.

Macknik, S. L., & Martinez-Conde, S. (2008, October/November). A perspective on 3-D visual illusions. *Scientific American Mind, 19*, 20–23.

Martinez-Conde, S., & Macknik, S. L. (2007). Mind tricks. *Nature, 448,* 414.

Martinez-Conde, S., & Macknik, S. L. (2008). Magic and the brain. *Scientific American, 299,* 72–79.

Pohl, R. F. (Ed.). (2004). *Cognitive illusions.* East Sussex, UK: Psychology Press.

Purves, D., Wojtach, W. T., & Howe, C. (2008). Visual illusions: An empirical explanation. *Scholarpedia, 3*(6), 3706.

Schrauf, M., Lingelbach, B., & Wist, E. R. (1997). The scintillating grid illusion. *Vision Research, 37,* 1033–1038.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28,* 1059–1074.

Simons, D. J., Lleras, A., Martinez-Conde, S., Slichter, D., Caddigan, E., & Nevarez, G. (2006). Induced visual fading of complex images. *Journal of Vision, 6*(10), 1093–1101.

Todrank, J., & Bartoshuk, L. M. (1991). A taste illusion: Taste sensation localized by touch. *Physiology & Behavior, 50,* 1027–1031.

Tse, P. U. (2005). Voluntary attention modulates the brightness of overlapping transparent surfaces. *Vision Research, 45,* 1095–1098.

O

Object Perception

The environment is full of elementary features, such as shapes, colors, and textures. However, observers do not perceive these elements in isolation. Rather, they combine them into two-dimensional objects, such as red circles, or into three-dimensional objects, such as cats and people. This entry discusses *object perception,* broadly defined as the ability to combine elementary features into whole objects. The ability to perceive objects is an important precursor to recognize three-dimensional objects in the environment. Recognition can occur at different levels of specificity: Observers can identify an object as one they have seen before (e.g., their neighbor's pet cat), or they can categorize that object as belonging to a more general class (e.g., as an animal rather than a vehicle). This recognition process occurs very quickly and accurately, which allows observers to successfully interact with objects in their environment.

The basic problem of object perception is the changing nature of the visual input. Light is reflected from the surface of objects and picked up by millions of photoreceptors on the retinal layer at the back of each eye. Thus, the retina creates a two-dimensional image of the three-dimensional world. This retinal image is the input to the visual system. The input changes because observers and objects move relative to each other or lighting conditions vary. Features are extracted from this changing retinal image and subsequently used to perceive and recognize objects. Psychologists, neuroscientists, and computer scientists have all tackled this problem in perception. Consequently, different approaches have emerged to study object perception, which have ultimately led to different theories of object recognition.

Structuralism and Gestalt Psychology

One of the earliest ideas about perception arose in the laboratory of Wilhelm Wundt, who established the first experimental psychology laboratory in 1879. Wundt believed that there were sensations evoked by physical stimulation of the corresponding sense, such as the sensation of heat, sound tone, or light intensity. He proposed that perception emerged from the combination of these pure sensations, and his students trained themselves to systematically isolate and measure the quality and intensity of these sensations through a technique of introspection or self-observation. This approach was later called structuralism by Wundt's student, Edward Titchener.

However, Max Wertheimer, Wolfgang Köhler, and Kurt Koffka found percepts that had no corresponding pure sensation, contrary to the tenets of structuralism. For example, Wertheimer found that briefly flashing two bars of light a short distance apart, one after the other, gave rise to the percept of a single bar moving from one location to another (which he called apparent motion or phi phenomenon). In apparent motion, there is no stimulus between the two bars to physically stimulate the retina and evoke a pure sensation of light. Consequently, Wertheimer and his colleagues founded an approach called Gestalt psychology,

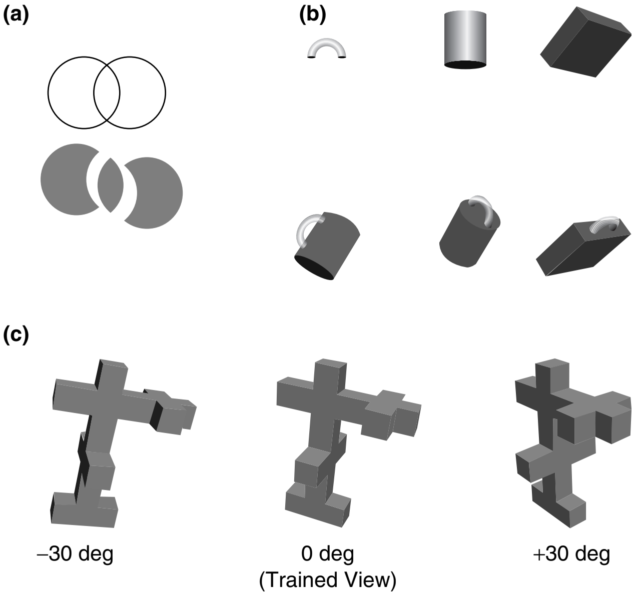


Figure 1 Objects Used in Object-Perception Theory and Research

Notes: (a) An ambiguous figure to illustrate the law of Prägnanz. (b) Three of the 36 geons proposed by Biederman. The combination of these geons and their spatial relationship form different objects. (c) An example of Tarr and Pinker's novel object seen from different viewpoints.

which focused on characterizing percepts rather than sensations, and formulating the laws that govern their creation. These laws became the Gestalt principles of perceptual organization. One of the central laws is the law of Prägnanz, roughly meaning "good figure," which states that a stimulus is perceived as simply as possible. Consider the stimulus illustrated on the top of Figure 1(a). This figure is ambiguous because it can have more than one perceptual interpretation. It can be interpreted as two interlocking rings. A more complex, but equally valid, percept is of three shapes, shown in gray, abutting each other. Observers mostly report the first percept, which is simpler. Other laws include: proximity—elements that are close together are grouped together as a percept; and similarity—elements that are similar to each other (e.g., they have the same color) are grouped together. These laws help group elements that might otherwise be ambiguous into whole objects.

Modern Approaches to Object Perception

In his influential book, David Marr popularized a computational approach to object perception and recognition. For Marr, the ability to perceive and recognize objects is essentially a problem in information processing. He conceptualized a series of stages. At each stage, information is transformed into a different representation and transmitted to the next stage. His main stages included a primal sketch, a two-and-a-half dimensional (2.5-D) sketch, and finally a three-dimensional (3-D) model. The primal sketch transforms light intensities on the retinas into edges and regions. The 2.5-D sketch then groups these edges and regions into visible surfaces. This representation encodes information about color, texture, and even distance to the observer at each point on the surface. The representation at this stage is 2.5-D because of the additional depth information that is encoded. Finally, a full 3-D mental model of an object is reconstructed from the 2.5-D sketch. This model specifies object parts relative to each other (e.g., the hand is connected to the wrist, the wrist is connected to the arm, the arm is connected to the shoulder, and so on). This spatial coordinate system is called an object-centered reference frame. Marr's computational approach helped link the Gestalt psychologists' various principles of perceptual organization to other visual functions, such as grouping parts into 3-D models. He also popularized the idea that object perception is about reconstructing the 3-D scene.

Another modern approach that built on the early ideas of object perception is Anne Treisman's feature integration theory. She suggested that attention is needed to group elements into percepts, and proposed that this grouping process occurred in two stages. First, there is a preattentive stage, in which elementary features, such as color and line orientation at different positions on the retinas, are automatically analyzed. Then, in a focused attention stage, these features are combined into objects. This combination process requires attention to integrate features belonging to an object. Without attention, the visual system sometimes incorrectly combines features from different objects leading to illusory conjunctions. For example, if a red "X" is flashed with a blue "K," participants may report seeing a red "K" if they are not attentive.

Theories of Object Recognition

The stable percept of an object is an important step toward recognizing that object, but one challenge is the changing visual input. Therefore, different theories of object recognition have been proposed to explain how observers solve what is known as the invariance problem. This problem states that observers perceive the same object, and can identify or categorize it, despite possible differences in the retinal images of that object. For example, when observers see an object from two different viewpoints, the 2-D image created on the retina is different between those two views. Similarly, different lighting conditions create different retinal images of the same object. Researchers agree that the underlying representation that supports this kind of robust recognition is critical; however, the nature of this representation remains a key issue in object recognition. Thus, theories of object recognition generally fall into two broad classes that differ on the features used for the object representation.

Structural Description Theories

Proponents of structural description theories propose that objects are represented by parts and their spatial relationships, which together form a structural description of an object. These descriptions discard an object's color and texture, for example, as the appearance of surface properties change with changes in viewing conditions (e.g., a change in lighting can change how color appears to an observer). The basic idea is that the same structural description can be recovered or otherwise derived from different retinal images of the same object. This robustness remains an appealing aspect of structural description theories despite the loss of surface information. Structural description theories have also been referred to as part-based or edge-based theories, given their reliance on parts and edges.

The first viable structural description theory for human object perception was proposed by David Marr and Keith Nishihara. According to their theory, object parts (e.g., a cat's leg) are represented by 3-D primitives called generalized cones, which specified arbitrary 3-D shapes with a set of parameters. For example, a cylinder can be produced by taking a circular cross section and sweeping it along a straight line. The circle traces out a cylinder with the line forming the main axis of that cylinder. By comparison, a rectangular cross section sweeps out the surface of a brick. More complex 3-D shapes can similarly be produced by sweeping different 2-D cross sections across different axes.

One of the challenges faced by Marr and Nishihara was how 3-D generalized cones can be recovered from 2-D images. They suggested that an object's bounding contour—the outline of an object in a picture—could be used to find the axes of its main parts. These axes could then be used to derive generalized cones and their spatial configuration. Recognition could then proceed by matching the structural description recovered from the image to those stored in visual memory. Thus, Marr and Nishihara try to solve the invariance problem by recovering view-invariant 3-D models from images.

Following Marr and Nishihara's seminal 1978 work, Irving Biederman proposed another influential structural description theory in the mid-1980s—recognition by components (RBC). Biederman argues that objects are mentally represented by a set of 36 components and their spatial relationship. He called these geons, for "geometrical ions." Geons are a subset of the generalized cones proposed by Marr and Nishihara, three of which are shown on the top of Figure 1(b). The combination of these geons into structural descriptions can be used to create familiar objects like a mug, a pail, or a briefcase, as shown in the bottom of Figure 1(b).

RBC theory builds on Marr and Nishihara's structural description theory in two innovative ways. First, unlike generalized cones, geons only differed qualitatively from each other. For example, a geon's axis can only be straight or curved, whereas generalized cones can, in principle, have any degree of curvature. Biederman's second innovation was to propose a more direct means to recover geons from images. According to RBC theory, geons are recovered from nonaccidental properties. These are properties of edges in an image (e.g., lines) that are associated with properties of edges in the world. To understand nonaccidental properties, consider seeing a box from many different viewpoints. From most views, observers see three sides of the box, which terminates in a "Y"-junction at a corner. This two-dimensional junction is an example of a nonaccidental property, and it is associated with a three-dimensional corner. From a few viewpoints, the Y-junction is not visible in the image such as when observers see one side of the box. It is therefore ambiguous whether such an image is a three-dimensional box or a two-dimensional rectangle. However, these “accidental” viewpoints are much less likely to be encountered relative to nonaccidental viewpoints. Thus, nonaccidental properties are highly robust (though not entirely invariant) to changes in viewpoint, viewing distance, and illumination.

Biederman suggested that two to three geons are enough to represent many objects. He also stated that RBC theory accounts for recognizing objects at what psychologists call the basic level. For example, “dog” is at the basic level, “poodle” is at a more specific subordinate level, and “animal” is at a more general superordinate level. To recognize objects at the subordinate level requires qualitatively different representations than the structural descriptions proposed by RBC theory. Biederman, for example, suggested that face recognition is based on the representation of fine metric details, such as the distance between eyes, rather than geons.

To solve the invariance problem, Biederman proposed a principle of componential recovery: Objects can be identified if their component geons can be identified. As geons are recovered from highly stable, nonaccidental properties, the same geon can be recovered from many different viewpoints, viewing distances, and illumination. Perhaps the strongest evidence for RBC theory comes from Biederman’s contour deletion studies. He and his colleagues took line drawings of everyday objects like a cup. They then deleted contours that could be used to recover geons (e.g., at junctions) or deleted the same amount of contours from other sections that could not be used to recover geons. They found that observers had no problem naming line drawings with preserved junctions but had difficulty naming line drawings with deleted junctions.

Image-Based Theories

In contrast to structural description theories, proponents of image-based theories posit that objects are represented as measurements of features that preserve many aspects of an object as they appear to an observer. For example, the visual system measures features such as color, texture, and even shading patterns on a surface. It can also encode the spatial location of these features. These measurements constitute a view of an object, which depict its appearance under specific viewing conditions, much like a picture taken by a camera from a fixed viewpoint. If a viewing condition changes (e.g., the observer walks to a different location to view an object), these measurements also change. In effect, object perception is supported by representations that are copies of the retinal images. For example, whereas structural descriptions discard surface features (e.g., color), an image-based representation would encode all visible features. Image-based theories have also been called view-based and appearance-based theories.

The image-based approach emerged in the early 1990s as a result of two key developments. First, there was an accumulation of evidence that indicated that observers were highly sensitive to viewing conditions, contrary to the predictions of structural description theories. For example, Michael Tarr and Steven Pinker trained observers to recognize block-like objects, which they had never seen before, from a specific viewpoint. Tarr and Pinker then rotated these objects away from their familiar trained view and tested observers’ ability to recognize them. The effect of depth rotation on the appearance of these novel objects is illustrated in Figure 1(c). Tarr and Pinker found that observers responded more slowly and made more mistakes to novel views relative to trained views. Second, Tomaso Poggio, Shimon Edelman, and Heinrich Bülthoff developed computational models that learned to match novel images of an object to trained images of that object. Importantly, these models do not reconstruct an object’s 3-D structure. Rather, they store a collection of views, and match a novel image to a candidate stored image based on the similarity of the two images. In addition to providing a theoretically rich framework for understanding object recognition, the behaviors of these models were similar to the behaviors of humans.

Tarr and Pinker synthesized these developments to address the invariance problem in their multiple-views theory of object recognition. They posited the idea that a 3-D object can be represented as a linked collection of experienced 2-D views. To explain how novel views are recognized, they further postulated a time-consuming mental rotation process, analogous to physically rotating an object, to transform a novel view until it matched a stored view. Observers could mentally rotate the input image until it matched a stored view. Accordingly, the time to perform this operation and how accurately it was performed should be proportional to the difference in viewpoint between the input view and stored view. Thus, Tarr and Pinker predicted that performance should depend on the orientation difference between the novel view and the nearest view encoded in memory. This is what they found in their experiments. Poggio and his colleagues proposed another normalization mechanism called view interpolation. In this model, stored views are used to capture possible appearances of an object. The stored views allowed the model to interpolate, or estimate, novel views of an object for recognition purposes. The idea is analogous to fitting a smooth curve through known data points and then predicting unknown data from this curve.

Finally, according to image-based theories, parts also play a role in object perception. However, unlike structural description theories, these parts are fragments extracted from specific images of objects from a particular class (e.g., different cats), rather than being prespecified (e.g., as a small set of geons). An image has a very large number of fragments that can be of any size. However, only a very small proportion of the possible fragments are useful for object recognition. Shimon Ullman uses a statistic called mutual information to find these informative fragments. This measure gives the probability that an object is present in an image if a fragment is present in the image. For example, if a human eye is present in the image, then there is a good chance that a human face is also present in that image. In this way, mutual information maximizes the ability for a set of fragments to distinguish between different objects (e.g., cats from dogs and horses). The set of informative fragments for a class is related to regularities that exist across members of that class. For example, cats are more likely than humans to have pointy ears and whiskers. There is both behavioral and neural evidence that human observers use fragments in object perception.

Two Versus Three Dimensions

Structural description models are often conceptualized as 3-D models because the primitives can be explicitly three-dimensional, as in Marr and Nishihara's generalized cones. Likewise, image-based models are often conceptualized as 2-D models because of their analogy to a camera. However, neither class of theories needs to be strictly two- or three-dimensional. For instance, Biederman argued that geons capture 2-D nonaccidental properties, which do not drastically change with viewing conditions, rather than 3-D structure per se. Similarly, Tarr suggested that information about depth at points on a visible surface can be a feature that is measured by the visual system and used for object perception. Shading cues, for example, can provide this additional depth information. However, depth information can only be recovered from visible surfaces, so the representation is neither strictly two-dimensional nor strictly three-dimensional—akin to Marr's 2.5-D sketch. This issue of dimensionality points toward a common ground for structural description and image-based theories of object recognition.

Quoc C. Vuong

See also Bayesian Approach; Face Perception; Object Perception: Physiology; Perceptual Development: Object Perception; Perceptual Organization: Vision; Perceptual Segregation; Vision; Vision: Cognitive Influences; Visual Scene Statistics

Further Readings

Biederman, I. (1987). Recognition-by-Components: A theory of human image understanding. *Psychological Review, 94,* 115–147.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* San Francisco: W. H. Freeman.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London B, 200,* 269–294.

Peissig, J. J., & Tarr, M. J. (2007). Visual object recognition: Do we know more now than we did 20 years ago? *Annual Review of Psychology, 58,* 75–96.

Peterson, M. A. (2001). Object perception. In E. B. Goldstein (Ed.), *Blackwell handbook of perception* (chap. 6, pp. 168–203). Oxford, UK: Blackwell.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature, 343,* 263–266.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21,* 233–282.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Ullman, S. (2007). Object recognition and segmentation by a fragment-based hierarchy. *Trends in Cognitive Science, 11,* 58–64.

OBJECT PERCEPTION: PHYSIOLOGY

Object perception or object recognition is the process in which visual input is assigned a meaningful interpretation that is available to perceptual awareness. It is fundamental to our ability to interpret and act in the world. Object perception is thought to occur through computations across a hierarchy of processing stages in the visual cortex, named the *ventral visual pathway.* This pathway begins in the primary visual cortex, area V1 in the occipital lobe, and ascends to regions in the lateral occipital cortex and the ventral occipitotemporal cortex. Damage (such as lesions, stroke, disease) to higher level visual areas in this pathway leads to specific deficits in object perception, such as the inability to recognize objects (object agnosia) and/or inability to recognize faces (prosopagnosia), while not affecting other visual abilities, such as determining the motion of objects or their contrast. As such, higher level regions in the ventral stream are thought to be necessary for conscious object perception.

Before the advent of noninvasive neuroimaging methods, such as functional magnetic resonance imaging (fMRI) in the mid-1990s, knowledge about the function of the ventral stream was based on single unit electrophysiology measurements in monkeys in lesion studies. These studies showed that neurons in the monkey inferotemporal (IT) cortex respond selectively to shapes and complex objects, and that lesions to the ventral stream can produce specific deficits in object recognition. However, it is difficult to make inferences on the relation between responses of single neurons in the monkey's brain and its perception of objects. Further interpreting lesion data is complicated because lesions are typically diffuse, they usually disrupt both a cortical region and its connections, and they are not replicable across patients.

With the advent of neuroimaging there has been an enormous advancement of our understanding of the neural basis of object perception in humans. The first set of fMRI studies of object and face perception in humans identified the regions in the human brain that respond selectivity to objects and faces. Then, a series of studies demonstrated that activation in object-selective and face-selective regions correlates with success at recognizing objects and faces, respectively, providing striking evidence for the involvement of these regions in perception. Once researchers found which regions in the cortex are involved in object perception, the focus of research shifted to examining the nature of representations and computations that are implemented in these regions to understand how they enable efficient object recognition in humans. This entry covers object-selective regions in the human brain, the role of the lateral occipital complex in object perception, neural mechanisms of invariant object perception, and category-selective regions in the human ventral stream.

Object-Selective Regions in the Human Brain

Object-selective regions in the human brain consist of a constellation of regions in the lateral and ventral occipitotemporal cortex, called the lateral occipital complex (LOC). These regions are defined functionally. Using fMRI, researchers scan participants' brains while the participant views different kinds of visual stimuli, and thus can determine which parts of the brain respond selectively to certain stimuli over other stimuli. Researchers found a constellation of regions in the lateral occipital cortex that respond more strongly when subjects view pictures of objects (e.g., cars, abstract sculptures) than when they view scrambled images of these objects, textures, or patterns. The LOC includes a posterior region, LO, located adjacent and posterior to a region involved in processing visual motion (medial temporal, MT, area) and a ventral region, pFus/OTS, which overlaps the occipitotemporal suclus (OTS) and the posterior part of the fusiform (Fus) gyrus (see color insert, Figure 28).

LOC responds similarly to many kinds of objects and object categories, including novel objects. It is thought to be in the intermediate- or high-level stages of the ventral stream visual hierarchy after early visual areas (V1, V2, V3, and V4) because it responds to specific visual stimuli (objects and shapes; whereas early visual areas respond to all visual stimuli, including texture patterns) and shows lesser sensitivity to low-level visual information, such as stimulus contrast, stimulus position, and stimulus size, compared to early visual regions. There are also object-selective regions in the dorsal stream, which are located on the dorsal part of the occipital lobe, including a region lateral to V3a, and regions overlapping the intraparietal sulcus (IPS) in the parietal lobe. However, these dorsal object-selective regions do not correlate with object perception and may be involved in computations related to visually guided actions toward objects. A review of the dorsal object-selective cortex is beyond the scope of this entry.

The Role of the Lateral Occipital Complex in Object Perception

Although the LOC is activated strongly when subjects view pictures of objects, this does not by itself prove that it is the region in the brain that 'performs' object recognition. There are many differences between objects and scrambled objects (or objects and textures). Objects have a shape, surfaces and contours; they are associated with a meaning and semantic information; and they are generally more interesting than textures. Each of these factors may affect the higher LOC response to objects than control stimuli. Nevertheless, several studies show that activation in the LOC is correlated with object perception rather than low-level features in the visual stimulus.

Converging evidence from multiple studies revealed an important aspect of coding in the LOC: It responds to object shape, not low-level or local visual features. These studies used stimuli that were controlled for low-level visual information across objects and nonobject controls and showed that LOC responds only when there is a global shape in the stimulus even when the low-level information is identical. These experiments have shown that LOC responds more strongly to the following: (1) objects defined by luminance compared to nonobjects generated from luminance (including bars, patterns, and scrambled objects), (2) objects generated from random dot stereograms rather than formless random dot stereograms, (3) objects generated from structure-from-motion relative to random motion and (4) objects generated from textures rather than flat texture patterns made of the same basic components. Further, LOC's response to objects is also similar across object formats (e.g., gray-scale photographs, black and white line drawings of objects, and object silhouettes) and it responds selectively to objects delineated by both real and illusory contours. Overall, these data suggest that LOC's activation to objects exhibits *cue invariance*. That is, it is insensitive to the specific visual cues that define an object.

Several recent studies also showed that the LOC responds to shape rather than surfaces or local contours. LOC's responses are higher to global shapes than surfaces or planes, and higher to global shapes than scrambled contours of the same shapes. For example, occluding bars presented in front of an object (which resembles the real life situation of seeing objects behind blinds) or stimuli with incomplete contours that still define a clear global shape (e.g., drawing an object's contour in which some of the strokes are missing), do not significantly reduce LOC's response as compared to unoccluded shapes or shapes with complete contours. These data underscore the findings that LOC's responses are driven by shape rather than low-level visual information that generates form or local features in the visual stimulus.

Importantly, LOC activations are correlated with subjects' object recognition performance. High LOC responses correlate with successful object recognition, and low LOC responses correlate with trials in which objects are present, but are not recognized. The correlation between object perception and brain activation in the LOC has been demonstrated using various methods. Some researchers have monotonically varied a parameter that affects the ability of subjects to perceive objects and have tracked the correlation between changes in brain activation and subjects' recognition performance. Parameters that have been manipulated include image presentation duration, the amount of visible object behind occluding blinds, object contrast, and the degree of coherence of objects presented in noise. Such studies have shown that pre-recognition activation in the LOC is lower than when objects are recognized, and that the level of activation increases monotonically until objects are recognized. Other studies used a different approach in which they showed subjects pictures close to *recognition threshold* (the border between seeing and not seeing), and compared brain activation in trials where subjects recognized objects to brain activation in trials where objects were present but not recognized. These experiments showed that LOC responses are higher in trials in which subjects detected objects than when they did not detect objects and LOC's responses were highest when objects were correctly identified.

Neural Mechanisms of Invariant Object Perception

One of the reasons that object perception is difficult is because there is significant variability in the appearance of an object (e.g., various views of a car look very different) which necessitates considerable generalization, yet the visual system needs to retain specificity to discriminate between objects that are similar (e.g., different cars contain the same parts arranged in a similar configuration). There are many factors that can affect the appearance of objects including the following: (a) the object being at different locations relative to the observer, (b) the two-dimensional projection of a three-dimensional object on the retina varies considerably due to changes in its viewpoint relative to the observer, (c) differential illumination which affect objects' color, contrast, and shadowing, and (d) occlusion. Nevertheless, humans are able to recognize objects in a tenth of a second across large changes in their appearance. This ability is referred to as *invariant object recognition.*

How does the LOC deal with the variability in objects' appearance? One view suggests that invariant object recognition is accomplished because the underlying neural representations are invariant to the appearance of objects. Thus, there will be similar neural responses even when the appearance of an object changes considerably. Other theories suggest that invariance may be generated through a sequence of computations across a hierarchically organized processing stream in which the level of invariance increases from one level of the processing to the next. For example, at the lowest level neurons code local features and in higher levels of the processing stream neurons respond to more complex shapes and are less sensitive to object transformations.

The fMRI studies of invariant object perception found differential sensitivity across the ventral stream to object transformations such as size, position, illumination, and viewpoint. Intermediate regions such as LO show higher sensitivity to image transformations than higher-level regions such as pFus/OTS. That is, LO shows sensitivity to object size, position, illumination, and rotation, whereas pFUS/OTS shows lesser sensitivity to object size and position, but still shows considerable sensitivity to object view and illumination. Notably, accumulating evidence from many studies suggests that at no point in the ventral stream are neural representations entirely invariant to object transformations. These results support an account in which invariant recognition is supported by a pooled response across neural populations that are sensitive to object transformations. One way in which this can be accomplished is by a neural code that contains independent sensitivity to object identity and transformation.

One object transformation that has been extensively studied is sensitivity to object position. Position invariance is thought to be accomplished in part by an increase in the size of neural receptive fields along the visual hierarchy. That is, as one ascends the visual hierarchy, neurons respond to stimuli across a larger part of the visual field. Findings from monkey electrophysiology suggest that even though neurons at the highest stages of the visual hierarchy respond to stimuli across a large part of the visual field, these neurons retain some sensitivity to object position and size. In humans, both the LOC and category-selective regions in the ventral stream (see below) respond to objects presented at multiple positions and sizes. However, the amplitude of response to object varies across different retinal positions. LO, pFUS/OTS, as well as category-selective regions respond more strongly to objects presented in the contralateral versus ipsilateral visual field. Some regions, such as LO, also respond more strongly to objects presented in the lower visual field than upper visual field. Responses also vary with eccentricity: LO, fusiform face-selective regions, and word-selective regions in the occipitotemporal sulcus respond more strongly to centrally than peripherally presented stimuli. In contrast, place-selective regions respond more strongly to peripherally than centrally presented stimuli.

These studies of sensitivity to object position indicate that neural populations in the human ventral stream retain some degree of sensitivity to position. How then is position invariant recognition accomplished if position information is present in the neural response? Current theories suggest that this is accomplished by maintaining separable information about object identity and object position across the response of a neural population that codes both identity and position. First, maintaining separable information about object position and identity may allow preserving information about the structural relationships between object parts (e.g., keeping information such as "the head is above the torso" and "the torso is above the legs"). Second, it may provide a robust way for generating position invariance by using a population code. That is, the pooled response across a population of neurons may generate position invariant perceptions. For example, consider a population of neurons. Each neuron's response is modulated (that is, its response varies) according to the position of the object in the visual field. However, at each given position in the visual field each neuron's response is higher for one object over other objects. Since the higher response for an object is consistent across positions it is possible to determine the object identity independent of position. Third, separable object and position information may allow concurrent localization and recognition of objects. That is, recognizing both *what the object is* and also determining *where it is*. This possibility challenges the prevailing hypothesis that position and identity information are represented in two parallel processing streams in which object identity ("what") information is thought to be represented in the ventral processing stream and object position ("where") is thought to be represented in the dorsal processing stream.

Category-Selective Regions in the Human Ventral Stream

In addition to the LOC, researchers found several additional regions in the ventral stream that show preferential responses to specific object categories (see color insert, Figure 28). Searching for regions with categorical preference was motivated by reports that suggested that lesions to the ventral stream can produce very specific deficits—such as the inability to recognize faces, or the inability to read words, whereas other visual (and recognition) faculties are preserved. By contrasting activations to different kinds of objects, researchers found ventral regions that show higher responses to specific object categories, such as a region in the left occipitotemporal sulcus that responds more strongly to letters than textures (the "visual word form area," VWFA); several foci that respond more strongly to faces than other objects (including the fusiform face area, FFA); regions that respond more strongly to houses and places than faces and objects (including a region in the parahippocampal gyrus, the parahippocampal place area, PPA); and regions that respond more strongly to body parts than faces and objects (including a region near MT called the "extrastriate body area," EBA, and a region in the fusiform gyrus, the "fusiform body area," FBA). Nevertheless, many of these object-selective and category-selective regions respond to more than one object category.

Findings of category-selective regions initiated a fierce debate about the principles of functional organization in the ventral stream. Are there regions in the cortex that are specialized for any object category? Is there something special about computations relevant to specific categories that generate specialized cortical regions for these computations? In explaining the pattern of functional selectivity in the ventral stream, four prominent views have emerged.

Limited Category-Specific Modules and a General Area for All Other Objects

Nancy Kanwisher and coworkers suggested that the ventral temporal cortex contains a limited number of modules specialized for the recognition of special object categories, such as faces (in the FFA), places (in the PPA), and body parts (in the EBA and FBA). The remaining object-selective cortex (LOC) is a general-purpose mechanism for perceiving any kind of visually presented object or shape. The underlying hypothesis is that there are few "domain-specific modules" that perform computations specific to these classes of stimuli beyond what would be required from a general object recognition system. For example, faces, like other objects, need to be recognized across variations in their appearance (a domain-general process).

However, given the importance of face processing for social interactions, there are aspects of face processing that are unique. Specialized face processing may include identifying faces at the individual level (e.g., John vs. Harry) and extracting gender information, gaze, and expression. These unique face-related computations may be implemented in specialized face-selective regions.

Process Maps

Michael Tarr and Isabel Gauthier proposed that object representations are clustered according to the type of processing that is required, rather than according to their visual attributes. It is possible that different levels of processing may require dedicated computations that are performed in localized cortical regions. For example, faces are usually recognized at the individual level (e.g., "That is Bob Jacobs"), but many objects are typically recognized at the category level (e.g., "That is a horse"). However, expert recognizers particularly excel at recognizing specific exemplars of their expert category (e.g., bird experts can easily distinguish between a "willow sparrow" and a "house sparrow"). These researchers hypothesized that expert processing uses processing similar to face processing because it requires fined-grained visual discriminations between items that have similar parts and configuration. Isabel Gauthier, Michael Tarr, and colleagues further showed that the FFA responds more strongly in birds and car experts when they viewed images of their expert category compared to other common objects. Therefore, they suggested that the FFA is involved in fine-grain discrimination between exemplars of any object category, and this processing is automatically recruited in experts.

Distributed Object-Form Topography

James Haxby and coworkers posited an "object form topography," in which the occipitotemporal cortex contains a topographically organized representation of shape attributes. The representation of an object is reflected by a distinct pattern of response across all ventral cortex, and this distributed activation produces the visual perception. Haxby and colleagues showed that the response to a given category could be determined by the distributed pattern of activation across all ventral-temporal cortex. Further, they showed that it is possible to predict what object category subjects viewed, even when regions that show maximal activation to a particular category were excluded. Thus, this model suggests that the ventral temporal cortex represents object category information in an overlapping and distributed fashion. One of the appealing aspects of this distributed representation is that it allows representation of a large number of object categories. Second, this model posits that both weak and strong signals in the ventral processing stream convey useful information about object category.

Topographic Representation

Rafael Malach and colleagues suggested that eccentricity biases underlie the organization of ventral regions because they found a correlation between category preference (higher response to one category over others) and eccentricity bias (higher response to a specific eccentricity than other eccentricities). They showed that regions that prefer houses to objects also respond more strongly to peripheral than central (foveal) visual stimulation. In contrast, regions that prefer faces or letters respond more strongly to centrally presented stimuli than peripherally presented stimuli. Rafael Malach and colleagues proposed that the correlation between category selectivity and eccentricity bias is driven by spatial resolution needs. Thus, objects whose recognition depends on analysis of fine details are associated with central visual field representations, and objects whose recognition requires large-scale integration are associated with peripheral visual field representations.

Presently, there is no consensus in the field about which account best explains ventral stream functional organization. Much of the debate centers on the degree to which object proceeding is constrained to discrete modules or involves distributed computations across large stretches of the ventral stream. The debate is both about the spatial scale on which computations for object recognition occur and about the fundamental principles that underlie specialization in the ventral stream.

The research described in this entry shows how neuroimaging and neurophysiological research in the past decade has advanced our understanding of object representations in the human brain. These studies have identified¶ the functional organization of the human ventral stream, showed the involvement of ventral stream regions in object recognition, and have laid fundamental stepping stones in understanding the neural mechanisms underlying invariant object recognition.

Kalanit Grill-Spector

See also Cortical Organization; Face Perception; Perceptual Development: Visual Object Permanence and Identity; Perceptual Expertise; Recognition; Vision: Cognitive Influences; Visual Processing: Extrastriate Cortex; Word Recognition

Further Readings

Andresen, D. R., Vinberg, J., & Grill-Spector, K. (in press). The representation of object viewpoint in the human visual cortex (2008, Nov. 25). *NeuroImage.* (Epub ahead of print)

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191–197.

Golarai, G., Ghahremani, D. G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J. L., Gabrieli, J. D., et al. (2007). Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nature Neuroscience, 10*(4), 512–528.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron, 24*(1), 187–203.

Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience, 3*(8), 837–843.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293*(5539), 2425–2430.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302–4311.

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., et al. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron, 60,* 1126–1141.

Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience, 4*(5), 533–539.

Li, N., & DiCarlo, J. J. (2008). Unsupervised natural experience rapidly alters invariant object representation in visual cortex. *Science, 321*(5895), 1502–1507.

Parr, L. A., Hecht, E., Barks, S. K., Preuss, T. M., & Votaw, J. R. (in press). Face processing in the chimpanzee brain (2008, Dec. 17). *Current Biology.* (Epub ahead of print)

Sayres, R., & Grill-Spector, K. (2008). Relating retinotopic and object-selective responses in human lateral occipital cortex. *Journal of Neurophysiology, 100*(1), 249–267.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science, 311*(5761), 670–674.

Vinberg, J., & Grill-Spector, K. (2008). Representation of shapes, edges, and surfaces across multiple cues in the human visual cortex. *Journal of Neurophysiology, 99*(3), 1380–1393.

OBJECT PERSISTENCE

Suppose that while playing tennis, an unfortunate swing propels your tennis ball out of the court and into some bushes. When you go to retrieve it, you find two tennis balls there. Which is yours? This is a type of *correspondence problem:* You must determine which of the two tennis balls corresponds to the one that you just hit out of the court. Though we are seldom explicitly aware of it, the visual system faces this type of problem thousands of times per day, whenever we encounter an object. On every such encounter, the visual system must determine whether a current bit of visual stimulation reflects a new object in the field of view or an object that was already encountered a moment ago. This is the challenge of *object persistence:* the perception of the world not only in terms of discrete objects, but in terms of objects that retain their identities as the same individuals over time and motion. This entry notes how the problem of object persistence features in several aspects of perception, and it describes some of the ways that persistence is determined in visual processing.

The computation of object persistence in perception seems to happen instantly and effortlessly, but like many aspects of perception (from perceiving depth to perceiving faces), a great deal of processing must occur to arrive at the "effortless" percept of persisting objects. If the visual system did not solve this type of problem so frequently and so efficiently, though, visual experience would be incoherent: We would be able to perceive objects, but it would be as if the world was created from scratch at every instant. On a street corner, for example, you might see a car *there,* then a car *there,* then a car *there*—with no recognition that you are seeing the *same* car at different times and places, as it moves down the street. Complicating things further, an object might only be re-encountered after a brief pause, due to a blink or to a period of visual occlusion, and when the object reappears, it might be in a different location or have changed some of its visual features.

Three Examples

Object persistence plays a role in many types of visual processing. Three examples are described here.

Apparent Motion

Object persistence is a crucial feature of even one of the simplest possible visual stimuli: two flashes of light, seen at different times and different places. If the flashes are far enough apart in time or space, they will be seen as two independent visual events. If they appear nearby and in quick succession, though, then the visual system interprets the flashes as a single event, and we see a single object that seems to move quickly from one location to the other. This phenomenon of apparent motion, in essence, reflects perception making an "unconscious inference" about persistence (in an automatic and hardwired way): The visual system decides under such circumstances that the second flash was the same object as the first flash, and so constructs a percept of motion to link them.

The Tunnel Effect

A similar situation arises in more natural events, as when a moving object disappears behind an occluder, and then an object emerges from the other side. In this situation, is the second object the same individual as the first object? In fact, if the movements are smooth and natural enough, we will see the event in terms of the uninterrupted and uniform motion of a single object behind the occluder. This is not a decision that is made consciously, but it is rather a type of automatic inference that is made in visual processing. Indeed, in this type of tunnel effect, we sometimes irresistibly see a single moving object even when we know for certain that the display contained two separate objects. In fact, the tunnel effect is equally compelling even when the two objects have entirely different surface features (e.g., colors and shapes): In this case, we see a single object that changes its features while occluded, rather than two separate objects, one of which must have been initially "hiding" behind the occluder.

Multiple Object Tracking

When trying to cross a street, you may have to attend carefully in order to keep track of how and where several cars are moving. When playing basketball, you may have to keep track of how and where several teammates and opponents are moving. These situations have been studied in attention research in the multiple object tracking (MOT) task, created by Zenon Pylyshyn. Observers initially see a number of identical looking objects. A subset of these are then flashed to indicate their status as targets, after which all of the (again identical) objects begin moving independently and unpredictably about the display. When they stop moving, observers must indicate which of the objects are the original targets. Observers are able to succeed at MOT when tracking up to three or four objects, but tracking more is nearly impossible for most people. This ability reveals how sustained object-based attention may give rise to a type of object persistence, and the constraints on MOT reveal that we are only able to perceive a small number of persisting objects at a time.

Object Files

Object persistence is thought to be determined in online perception at the stage that is often called "midlevel vision." At this level, objects may be represented as the same individuals over time, despite changes to both their lower-level visual features (e.g., red, round) and their higher-level category descriptions (e.g., "apple"). The challenge of object persistence can be considered in terms of these kinds of mental representations: The visual system must decide whether each bit of stimulation reflects an object that was already encountered (which might occasion the updating of an existing object representation), or a new one (which might occasion the creation of a new object representation).

The most influential theory of such representations is the *object file* framework put forth by Anne Treisman and Daniel Kahneman. An object file is a visual representation that "sticks" to a moving object over time on the basis of how and where that object moves, and stores (and updates) information about what that object looks like. In this way, object files help to construct our conscious perception of objects, telling us "which went where," and underlying the perception of object persistence despite changes in either what an object looks like or momentary periods of occlusion. This framework can be used to interpret the examples noted earlier in terms of underlying representations. We see apparent motion because the two flashes are represented by the same object file over time, rather than by two separate object files. Similarly, we see the tunnel effect because the two objects are represented by the same object file. And we are able to track multiple objects in the MOT task because we are able to keep the same object files assigned to the target objects as they move.

Beyond these types of perceptual phenomena, the operation of object files can also be measured in terms of an object-specific priming effect (where priming is the enhancement of the speed or accuracy of response to a stimulus due to prior exposure to a similar stimulus). Suppose that two objects are initially presented in a display, and distinct symbols (perhaps letters) appear briefly on each of them. The objects then move about the display for a brief period, after which a single "probe" symbol appears on just one of the objects. A subject's task, in one variant of this type of experiment, is to press a key to indicate whether the final probe symbol was the same as any of the initially presented symbols. Subjects will be faster to indicate "yes" than to indicate "no" due to a type of display-wide priming. But their "yes" responses will be even faster when the probe is the same symbol that initially appeared on that same object, compared to when it was the symbol that had initially appeared on the other object. This difference in response time is termed an *object-specific preview benefit* (OSPB). Thus, this effect can be used as an index of object persistence: Manipulations that degrade object files will result in attenuated OSPBs.

Principles of Persistence

Researchers have made great strides in identifying some of the principles used by the visual system in order to perceive persisting objects. Three of the most salient principles are described here.

Continuity

A basic fact about the world is that objects cannot simply go in and out of existence over time. For two objects encountered at different locations to be subsequent stages of the same individual, there must be a spatiotemporally continuous path between them. If an object disappears at one location, and an object immediately appears at a different (spatially separated) location, then those two instances cannot be the same object. This constraint of continuity is used in visual processing (taking occlusion into account) to constrain whether two bits of stimulation are seen as the same object over time, and its role can be seen in each of the examples previously noted. We only see apparent motion, for example, when the two flashes occur close enough in space and time (in which case the spatiotemporal proximity is judged to reflect a single swiftly moving object, instead of a coincidence). Similarly, we only see the tunnel effect when the two objects are linked by a smooth spatiotemporal trajectory; and we are only able to track multiple objects that move along what appear to be spatiotemporally continuous trajectories. The use of continuity to perceive persisting objects also comes online at an early age: Even young infants appear to appreciate that objects must trace continuous paths (as evidenced in experiments that measure their looking times), and continuity is considered by some developmental researchers to be one of the principles of "core knowledge" that help drive further learning about the visual world.

Cohesion

Another principle that determines object persistence is that of cohesion: An object must maintain a single bounded contour in order to be represented as persisting over time. This is also an intuitive defining characteristic of what it means to be an object in the first place: If you attempt to grab an object and only part of it comes away in your hand, then it must not have been a single object. The operation of cohesion as a constraint on object persistence can be seen in several ways. For example, multiple object tracking becomes very difficult if the objects fail to maintain unitary boundaries over time, and violations of cohesion such as an object splitting into two greatly attenuate the resulting OSPBs that are observed in object-specific priming studies. Cohesion is also thought by some to be a principle of "core knowledge," and developmental researchers have shown that cohesion violations also cause young infants to lose track of objects.

Spatiotemporal Priority

In everyday life, we identify some objects as the same individuals over time not because of how they move, but because of what they look like: When you next see your best friends, for example, you will likely identify them not because you watched them trace a continuous path since your last encounter, but because they look like your best friends. This way of assessing persistence breaks down, though, when an object (such as your missing tennis ball) might look identical to other objects. Moreover, these two sorts of principles may come into conflict: Suppose, for example, that an object traces a continuous trajectory but suddenly changes what it looks like. On such occasions, visual processing seems to apply a principle of spatiotemporal priority: When deciding whether an object is the same persisting individual from some earlier time, factors relating to how and where that object has moved will almost always trump factors relating to what the object looks like. For example, we may automatically see apparent motion between two flashes even if those flashes are very different colors and shapes, so long as they occur near enough in space and time. In such cases, we may thus see a red circle transform into a green square during its movement, rather than seeing two separate objects. Similarly, if the spatiotemporal trajectory linking two objects in the tunnel effect is fast and smooth enough, we may irresistibly see a single object (which changed its features while occluded) rather than seeing two separate objects (one pre-occlusion and one post-occlusion) with different properties. And, multiple object tracking is largely unaffected by changing surface features, so long as the objects move continuously. This spatiotemporal priority also appears to be a primitive principle of object persistence, because even nonhuman primates experience the tunnel effect in such situations. And spatiotemporal priority may be explained in terms of object files. Spatiotemporal factors are thought to control the construction and maintenance of such representations, but surface features are just along for the ride: When they change, the only result may be that the contents of the object file are modified.

The Importance of Persistence for Other Aspects of Perception

The importance of object persistence for perception can be appreciated not only by considering how we perceive the world, but also in how such computations influence other aspects of visual processing. Some of these connections must exist almost by definition. For example, the perception of number requires object persistence, because you cannot perceive multiple distinct objects without being able to recognize whether two objects are the same or different individuals. Object persistence also influences less obvious aspects of perception, such as visual memory: Changes are more easily detected between two stimuli that are encoded as the same persisting object. And neuroscientific studies show that principles such as spatiotemporal continuity influence neural representations of objects as the same over time—even in regions such as the ventral cortex that are usually associated with the processing of what objects look like. Ultimately, the "currency" of many perceptual processes may be persisting visual object representations.

Brian J. Scholl and Jonathan I. Flombaum

See also Attention: Object-Based; Binding Problem; Event Perception; Motion Perception; Perceptual Development: Object Perception; Visual Memory

Further Readings

Dawson, M. (1991). The how and why of what went where in apparent motion: Modeling solutions to the motion correspondence problem. *Psychological Review, 98,* 569–603.

Flombaum, J. I., Kundey, S. M., Santos, L. R., & Scholl, B. J. (2004). Dynamic object individuation in rhesus macaques: A study of the tunnel effect. *Psychological Science, 15,* 795–800.

Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24,* 175–219.

Pylyshyn, Z., & Storm, R. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 179–197.

Ramachandran, V. S., & Anstis, S. M. (1986). The perception of apparent motion. *Scientific American, 245,* 102–109.

Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind & Language, 22,* 563–591.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology, 38,* 259–290.

Yi, D.-J., Turk-Browne, N. B., Flombaum, J. I., Kim, M., Scholl, B. J., & Chun, M. M. (2008). Spatiotemporal object continuity in human ventral visual cortex. *Proceedings of the National Academy of Sciences USA, 105,* 8840–8845.

OLFACTION

The sense of smell is one of nature's true wonders, capable of detecting and identifying thousands of diverse odorants at extremely low concentrations. Along with its sister sense of taste, this sensory system is the most ubiquitous in the animal kingdom, occurring in one form or another in nearly all air-, water-, and land-dwelling animals. Even single-celled organisms, such as amoebae and paramecia, have mechanisms for sensing environmental chemicals. In our own species, the sense of smell plays a significant role in daily life, largely determining the flavor of foods and beverages and providing an early warning for such environmental dangers as leaking natural gas, fire, toxic fumes, spoiled foods, and polluted environments. Indeed, all environmental nutrients and airborne chemicals required for life enter our bodies by the nose and mouth. In this entry, the olfactory receptors, bulbs, and cortex will be discussed, as well as odor perception and factors that influence olfaction.

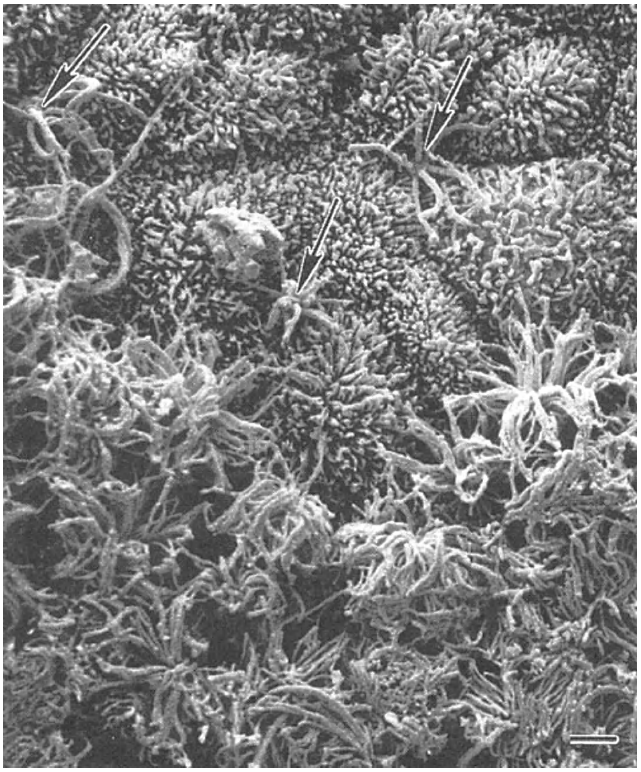


Figure 1 Transition Between Olfactory (Lower) and Respiratory (Upper) Epithelium

Source: From Menco & Morrison (2003). Copyright 2003 Marcel Dekker, Inc.

Note: Arrows indicate receptor cell endings with cilia. Bar = 5 μm (micrometers).

The Olfactory Receptors

An estimated 10 to 15% of the air that enters the nose upon inhalation reaches the olfactory epithelium, a membrane high in the nasal cavity that harbors specialized receptor cells that detect chemicals. Each of these 6 to 10 million receptor cells—the most receptor cells of any sensory system save vision—extends 10 to 30 threadlike cilia into the overlying mucus. Some cilia are quite long (e.g., up to 1/8th of a millimeter) and, in aggregate, provide an enormous surface area. In the human, this surface area is around six square inches, whereas in the German shepherd it is five times the surface area of the dog's body! Odorants contained within the inspired airstream absorb into the mucus and either diffuse to receptors located on the cilia or are carried through the mucus to such receptors by specialized carrier proteins. At some point, the residual odorants are deactivated by enzymes within the mucus and transported into the throat by the moving mucus blanket, where they are ultimately swallowed. A picture of the surface of the olfactory epithelium where it borders the respiratory epithelium is shown in Figure 1. Note the receptor cells with their multiple long cilia.

The 2004 Nobel Prize in Physiology or Medicine was awarded to Linda Buck and Richard Axel for identifying the gene family responsible for converting information coded by genes into the olfactory receptor cell proteins, that is, the olfactory receptors, located on the cilia of the olfactory receptor cells. In the human, around 450 types of such receptors are expressed, about half the number that is expressed in the mouse. This large number of receptor types contrasts significantly with the small number of receptor types found in vision; that is, three types of cones and one type of rod. It is currently believed that each of the 6 to 10 million receptor cells expresses only one type of receptor. However, each receptor can be activated by a number of different chemicals, and there is overlap between the responsiveness of receptors to the same chemical. For example, receptors a, b, c, d, j, n, w, and z may be activated by odorant A, whereas receptors c, d, j, n, q, and r may be activated by odorant B. Note that, in this example, receptor cells c, d, j, and n are activated by both odorants, leading to the concept that receptor cells are generalists.

The olfactory receptor cells have a number of unique features. First, they serve both as the receptor cell and the first order neuron. Second, their cilia lack the biochemical machinery found in other cilia responsible for motility, so the cilia do not beat in unison and more or less waft in the mucus. Third, when damaged they can be replaced by stem cells located deep in the epithelium, although in the human such replacement is rarely perfect and other cells types often replace regions previously occupied by the receptor cells. Fourth, these cells can be a direct conduit of exogenous agents, including viruses, bacteria, and toxins, from the nasal cavity into the brain. Indeed, it was known in the first half of the 20th century that polio virus commonly entered the brain via this route, leading to public health initiatives in Canada to chemically cauterize the olfactory region of school children in attempts to prevent the contraction of polio.

When enough odorant molecules bind to the receptors of a given cell, an ion flux occurs across the membrane and a propagating action potential is generated in the more elongated segment of the cell that projects into the olfactory bulb, termed its *axon.* This, in turn, results in the release of a chemical agent, generically termed a neurotransmitter, from the terminal ends of the cell into the space (i.e., the synaptic cleft) in which the cell communicates with cells in the olfactory bulb.

The Olfactory Bulbs

The paired olfactory bulbs are elongated ovid structures found underneath the brain roughly between the eyes in humans. These layered structures are not simple relay stations, but contain a wide range of interacting cell types that express many types of neurotransmitters. They are the sole recipients of the axons of the olfactory receptors, which first pass in bundles, termed the *olfactory fila,* from the nasal cavity into the brain via small holes in the cribriform plate of the ethmoid bone. The axons of the olfactory receptors then make up the outside layer of the olfactory bulb, termed the *olfactory cell layer.* The next layers of the bulb, in order, are the glomerular cell layer, the external plexiform layer, the mitral cell layer, the internal plexiform layer, and the innermost granule cell layer. In young people, the glomerular cell layer comprises thousands of 50- to 200-μm spherical *glomeruli*—the first relay stations of the olfactory system. The glomeruli are arranged in single or double layers. In older people, the glomerular layer becomes less defined and may even disappear subsequent to death of the olfactory receptor cells. The glomeruli are invariant components of both vertebrate and invertebrate olfactory systems.

Each olfactory receptor cell axon diverges from the bulb's olfactory nerve layer to penetrate a single olfactory glomerulus, where it multiplies branches and synapses with extensions, termed *dendrites,* of the next—second order—neurons in the system. The second order neurons are the mitral/tufted and periglomerular cells. A high-degree convergence (coming together) occurs between the incoming receptor cells and the second order cells; for example, in the rabbit only one second order neuron exits a glomerulus for every 1,000 primary receptor cells that enter it. This results in considerable summation of activity upon the second order neurons and is one reason for the exquisite sensitivity of the system.

The mitral/tufted cells send their axons via the olfactory tract to the olfactory cortex (see next section), whereas the periglomerular cells have localized projections confined to the bulb. The latter cells serve to modulate neural activity within the bulb itself, making connections, that is, synapses, with the cell bodies and dendrites of adjacent periglomerular cells, as well with dendrites of mitral and tufted cells. Such circuits are believed to sharpen the contrast between adjacent glomeruli and, in some cases, to recruit additional glomeruli as odorant concentration increases. Importantly, each glomerulus receives input from receptor cells that express the same type of receptor (see color insert, Figure 8). Interestingly, the type of receptor protein dictates the glomerulus to which the cell projects. Thus, in mice in which one olfactory receptor has been genetically substituted for another, the cell targets the glomerulus associated with the substituted receptor, not the original receptor.

Like the olfactory epithelium, some cells within the olfactory bulb are replaced by stem cells. Primordial cells germinate within a brain region located underneath the lateral ventricles (cavities within the brain that are continuous with the central canal of the spinal cord). From this subventricular region, they migrate to the olfactory bulb along a path known as the *rostral migratory stream,* ultimately repopulating periglomerular and granule cells.

The Olfactory Cortex

The mitral and tufted cell axons synapse on structures collectively termed the *primary olfactory cortex*—the anterior olfactory nucleus, the piriform cortex, regions of the amygdala and periamygdaloid complex, and the rostral entorhinal cortex. The left olfactory tract projects to the left side of the brain, and the right olfactory tract to the right side of the brain. Unlike other sensory systems, the olfactory system projects directly to the cortex, albeit a simpler three-layered cortex termed *allocortex,* without synapsing first in the thalamus, the major sensory relay system of the brain. This direct connection is believed by some to explain the strong associations between odors and memory, emotion, and endocrine function.

It is generally accepted that major elements of the olfactory cortex, such as the piriform and entorhinal cortices, are critical for odor identification and ultimately odor perception. Thus, the piriform cortex is active in tasks involving long-term odor recognition and the determination of odor familiarity. However, lesions within the primary olfactory cortex, such as those inflicted by ablating the amygdala and hippocampus for control of intractable epilepsy, influence a range of olfactory measures, including those of odor identification, detection threshold sensitivity, and odor discrimination.

Odor Perception

How do the elements of the olfactory system work together to produce the perception of an odor? It is important to realize that only rarely are smells denoted by single chemicals within the environment, yet our odor perceptions are clearly of a unitary nature. For example, the aroma of coffee depends upon hundreds of volatilized odorants, but we perceive only one unitary smell of coffee when smelling these vapors. Although trained coffee tasters, like wine tasters and perfumers, can detect a limited number of “notes” within the overall percept, the unitary sensation predominates.

What is the basis for such synthesis? It is generally believed that each odorant mixture activates a different combination of odor receptors, thereby inducing a unique spatial and temporal pattern of activity across the glomeruli. Some odorants within the mixture can block receptors that other odorants would ordinarily bind to, serving to filter the information that reaches the bulb. The pattern of activation observed across the glomeruli is further refined by the circuitry of the bulb. Hence, activation of the granule cells and periglomerular cells can influence the nature of the neural activity that is sent via the mitral and tufted cells to the olfactory cortex. Importantly, modulation of bulb activity is influenced by central structures that send information from the brain back into the bulb. Thus, the output signals of the bulb can be influenced by such factors as bodily state (hunger, arousal) and prior experience with the involved stimuli or similar sets of stimuli.

According to one theory, the primary olfactory cortex serves at least three key functions. First, when the pattern of transferred information is known (i.e., reflects a previously experienced odor), it is recognized within this cortex via a pattern matching system and is perceived as a familiar odor. Second, when this is not the case (i.e., the odor or a similar odor has not been previously experienced), the odor is perceived as strange and a new pattern is encoded into memory that will allow for future pattern matching activity. Third, combinations of chemicals can be recognized by the pattern matching system, even when some elements of the combinations are missing (i.e., the system can recognize an odorant in an environment with competing odors or in an environment that mutes some elements of the stimulus complex). The fact that key components of the olfactory cortex rapidly become insensitive (i.e., habituate) to non-novel odors has been suggested to increase the salience of novel odors.

This general process explains how we come to recognize odors, yet the involved brain structures also integrate, or take into account, information about the whole stimulus from which an odorant emanates. Hence, thinking about an odor or the visual presentation of a picture of an object associated with the odorant's source can evoke neural activity in the same brain regions activated by the odorant, showing that, in some cases, these brain regions "learn" the various elements of the stimulus. Both intra- and intermodal interactions have been reported in behavioral studies that are in accord with this concept. For example, in one study where a lemon-flavored candy was colored purple, most people reported the taste as being grape, not lemon. In another study subjects were first exposed to an unfamiliar odor mixture, smoky-cherry. Later, when exposed to the cherry odorant alone they reported it smelling slightly smoky and when exposed to the smoke odorant alone, they reported it smelling slightly cherrylike. This suggests what is being consciously perceived is, in large part, a memory code of the stimulus object, which goes beyond just its smell.

Factors That Influence the Ability to Smell

The ability to smell is not static; rather, it is influenced by such factors as age, sex, nutrition, health, accidents, smoking habits, and reproductive state. For instance, olfactory sensitivity for some odors is greater before than after a meal. In general, women outperform men on tests of odor discrimination, recognition, and detection, and retain normal smell function to a later age than men.

Smell loss is common in later life, with significant decrements occurring in approximately 1/2 of those between 65 and 80 years of age and 3/4 of those 80 years of age and older. Such dysfunction undoubtedly contributes to the fact that a disproportionate number of elderly people die in accidental natural gas poisonings and explains why many report that food has little flavor. In some cases, smell loss results in nutritional disturbances and even death. It is rarely appreciated that most food flavors (e.g., chocolate, meat sauce, strawberry, pizza) are dependent upon the stimulation of the olfactory receptors from the rear of the nose while chewing and swallowing. Taste buds primarily mediate such sensations as sweet, sour, bitter, and salty.

Another common cause of smell dysfunction is head trauma. Blows to the head that result in rapid acceleration or deceleration of the brain relative to the skull can sever or damage the very thin olfactory fila. A strange odor, likely representing either degeneration or attempts at regeneration of the receptor neurons, is often noticed for a few weeks or longer after such injuries. Olfactory dysfunction occurs more frequently from blows to the back of the head than to the front of the head, in part because frontal blows are cushioned to some degree by the collapse of soft facial structures, such as the nose and sinuses.

Of particular interest to the clinician is the fact that olfactory dysfunction is among the first signs of Alzheimer's disease and Parkinson's disease. Indeed, smell loss often occurs several years before the onset of the classical disease signs. Such findings have led to the use of olfactory tests in identifying people at the earliest stages of these diseases so that when neuroprotective treatments become available, they can be administered before significant irreversible neural damage has occurred.

Richard L. Doty

See also Common Chemical Sense (Chemesthesis); Flavor; Olfactory Adaptation; Olfactory Localization; Olfactory Quality; Olfactory Receptors and

Transduction; Olfactory Stimulus; Perceptual Development: Taste and Olfaction; Taste

Further Readings

Brewer, W. (Ed.). (2006). *Olfaction and the brain.* Cambridge, UK: Cambridge University Press.

Buck, L. B. (2005–2006). Unraveling smell. *Harvey Lectures, 101,* 117–134.

Doty, R. L. (2001). Olfaction. *Annual Review of Psychology, 52,* 423–452.

Doty, R. L. (Ed.). (2003). *Handbook of olfaction and gustation* (2nd ed.). New York: Marcel Dekker.

Doty, R. L. (2008). The olfactory vector hypothesis of neurodegenerative disease: Is it viable? *Annals of Neurology, 63,* 7–15.

Hudson, R. (1999). From molecule to mind: The role of experience in shaping olfactory function. *Journal of Comparative Physiology A, 185,* 297–304.

Johnson, B. A., & Leon, M. (2007). Chemotopic odorant coding in a mammalian olfactory system. *Journal of Comparative Neurology, 503,* 1–34.

Menco, B. P. O. M., & Morrison, E. E. (2003). Morphology of the mammalian olfactory epithelium: Form, fine structure, and pathology. In R. L. Doty (Ed.), *Handbook of olfaction and gustation* (2nd ed., pp. 17–49). New York: Marcel Dekker.

Mombaerts, P. (2006). Axonal wiring in the mouse olfactory system. *Annual Review of Cell and Developmental Biology, 22,* 713–737.

Morot, G., Brochet, F., & Dubourdieu, D. (2001). The color of odors. *Brain and Language, 79,* 309–320.

Pager, J. (1978). Ascending olfactory information and centrifugal influxes contributing to a nutritional modulation of the rat mitral cell responses. *Brain Research, 140,* 251–269.

Rozin, P. (1982). "Taste-smell confusions" and the duality of the olfactory sense. *Perception & Psychophysics, 31,* 397–401.

Stevenson, R. J., & Boakes, R. A. (2003). A mnemonic theory of odor perception. *Psycholological Review, 110,* 340–364.

Wilson, D. A., & Stevenson, R. J. (2006). *Learning to smell: Olfactory perception from neurobiology to behavior.* Baltimore: Johns Hopkins University Press.

OLFACTION: DISORDERS

We live in a sea of volatile molecules whose presence we are aware of largely or exclusively through our sense of smell. Smells warn us of environmental dangers (fire, toxic chemicals); they create a sense of intimacy with loved ones, whose signature odors we learn to recognize; they provide pleasure through perfumes and other desirable fragrances; and they constitute a substantial proportion of the flavors of food, as volatile molecules released in the oral cavity flow through the nasopharynx to the nose and stimulate olfactory receptors. Because changes in smell greatly disrupt the ability to appreciate food flavors, eating becomes a dissatisfying experience, and the many social occasions that center around food can become awkward and frustrating. Moreover, the inability to detect leaking gas or burning or spoiled food creates real dangers in daily life. Few people appreciate the range of information about our world provided by the sense of smell, or contemplate the impact of a disruption in that source of information. Yet, as demonstrated in this entry, olfaction is a vulnerable sense, and olfactory dysfunctions are common.

Vulnerabilities

Olfaction is subserved by a single cranial nerve located in a position that makes it particularly susceptible to damage. In traveling from the nose to the olfactory bulb (the first brain relay in the olfactory pathway), axons of the olfactory neurons pass through the cribriform plate of the ethmoid bone, which has a honeycomb-like appearance. As a result, they are subject to being crushed or severed when forces associated with head injury cause the brain to collide violently with the skull.

In addition, olfactory receptors are located in a small patch of tissue high in the convoluted structure of the nasal cavity. Any number of factors producing changes in nasal patency or airflow patterns can potentially limit the access of stimulus molecules to those receptors.

Finally, the olfactory neurons *are* the receptor cells. These neurons extend cilia on which smell receptors are located along the epithelial surface of the nasal cavity. Thus, they are uniquely exposed to the external environment and subject to a constant barrage of potentially toxic chemicals and particulates, as well as being susceptible to direct injury from microbes. Although olfactory receptor neurons are also highly unusual in that they are replaced on an ongoing basis throughout life, this is a complex process requiring reinnervation of the olfactory bulb, and it is often imperfect.

Types of Dysfunction

Most often, olfactory dysfunction takes the form of a simple loss of or diminution in sensitivity to smells. In general, *anosmia* is used to refer to the absence of smell function, and *hyposmia* to diminished smell sensitivity. However, *anosmia* does not necessarily imply a complete inability to detect the presence of a volatile odorous compound at any concentration, because most odorous compounds, at least at high concentrations, also stimulate nasal fibers of the trigeminal nerve and produce irritant and/or thermal sensations.

Individuals with olfactory dysfunction may also (often in combination with reduced smell sensitivity) experience an odor sensation in the absence of an odorous stimulus, or distortions in the perceived qualities of odorous stimuli. The terms *dysosmia* and *parosmia* are used interchangeably to refer to both of these conditions. In this entry, however, the former condition will be referred to more specifically as *phantosmia,* and the use of *dysosmia* will be limited to the latter.

Assessment

Several test batteries for the assessment of olfactory function in a clinical setting have been developed, and some are available commercially. All include a measure of threshold sensitivity (the lowest concentration of an odorant that can be detected) and a multiple-choice odor identification test using suprathreshold concentrations of odorants. Phenylethyl alcohol, an odor compound that has a roselike smell and elicits little or no nasal trigeminal response at any concentration, is widely employed as the threshold stimulus in clinical testing, but clinical norms for threshold sensitivity to butanol and pyridine have been established as well. The particular odors used in identification tests vary considerably, in part because they must be generally familiar to the individual being tested in order to provide a meaningful measure of olfactory function. What is a familiar smell in one culture may not be in another, and what is a familiar smell to adults may not be to children.

Variations in the stimuli presented and procedural details seem to have little impact on the results of olfactory testing in a clinical setting, because similar rates of diagnosis of olfactory dysfunction have been reported by all of the major chemosensory clinical centers. There are also high correlations between measures of threshold sensitivity and odor identification ability, suggesting that both types of test measure essentially the same property (sensitivity) and are redundant. However, this correlation breaks down in individuals complaining of dysosmia, reflecting the fact that most of these patients evidence reasonably good absolute sensitivity to the presence of odorous stimuli but have substantial difficulty identifying those stimuli, presumably due to the quality distortions they experience. Thus, the discrepancy between performance on threshold and identification measures may be diagnostically useful.

No performance measure has been found to distinguish phantosmia, so this diagnosis relies exclusively on patient report. Most patients who complain of odor phantoms do, however, evidence measurable olfactory dysfunction.

Etiologies

Modern chemosensory clinical centers have found diseases of the nose and/or paranasal sinuses to contribute to the plurality of cases of documented smell dysfunction. Smell losses associated with nasal/sinus disease (NSD) can be profound; indeed, the majority of NSD patients who present to chemosensory clinical centers are found to be anosmic. Reports of phantom odors, sometimes in association with odor quality distortions, are not uncommon in this patient group, but quality distortions by themselves are relatively rare.

Mechanical obstruction of the access of molecules to the olfactory receptors in rhinitis (and/or nasal polyposis) provides an easy explanation for this form of loss. However, it seems clear that pathologic mechanisms other than mechanical obstruction are also involved in NSD-related smell loss. For example, obstruction of the nasal airway in patients with allergic rhinitis has not been found to be associated significantly with olfactory sensitivity. Moreover, at least some patients with NSD and diminished olfactory function can be shown through endoscopic and computed tomography

(CT) scanning techniques to be free of significant obstruction of the olfactory area, but they nonetheless recover olfaction with treatment of their NSD. There is evidence of damage to the olfactory neuroepithelium secondary to chronic inflammation, which likely contributes to this form of loss; however, the fact that NSD-related losses often respond rapidly to the anti-inflammatory effects of systemic corticosteroid therapy suggests this is not the sole underlying mechanism. Alternatively, edema of the neuroepithelium could stretch the olfactory neurons and impede synaptic transmission. In addition, changes in the composition of the mucous overlying the olfactory receptor cilia could interfere with transport of odorant molecules to the receptors and/or with receptor binding. The possible role of such changes in NSD-related olfactory loss requires further study.

The second most common etiologic basis for smell dysfunction is prior upper respiratory infection (URI). Because viral URI can incite secondary bacterial sinusitis, patients with these two forms of smell dysfunction may present with similar histories. A number of characteristics of URI-related smell dysfunction do, however, clearly differentiate it from NSD-related loss.

First, URI-related dysfunctions tend to occur less frequently in young individuals than do NSD-related losses. The post-URI patient group also includes proportionately more women than does the NSD group. In addition, URI-related dysfunctions are significantly less likely to manifest as anosmia than are NSD-related dysfunctions. On the other hand, several authors have reported that dysosmia and phantosmia are significantly more frequent components of URI- than of NSD-related olfactory dysfunction, with approximately half of the post-URI patients reporting one or both of those symptoms.

Finally, damage to peripheral olfactory receptors appears to be the sole underlying mechanism for URI-related dysfunction. These patients do not respond to anti-inflammatory therapies, and histopathological studies of postviral olfactory dysfunction have demonstrated varying degrees of olfactory epithelial destruction in these patients, ranging from virtually total destruction of olfactory receptor neurons to reductions in the number of receptors, with patches of epithelium having a relatively normal appearance. Furthermore, the severity of histopathologic change has been found to be correlated with the observed olfactory deficit. Either abnormal axonal reconnection of regenerating neurons or the patchy degeneration (and/or incomplete regeneration) observed might underlie the odor quality distortions often reported by URI patients.

The third major etiologic basis for smell dysfunction is head trauma. The first reports of post-traumatic anosmia appeared in the medical literature in the latter half of the 19th century. Current estimates suggest that 20 to 30% of head trauma patients sustain some degree of olfactory impairment. The likelihood of posttraumatic olfactory loss appears to increase with the severity of the injury, although it can occur after trivial injuries with no associated post-traumatic amnesia, and blows to the occipital region may be most likely to produce smell dysfunction. As is the case in NSD-related loss, the majority of trauma patients who present to chemosensory clinical centers are found to be anosmic, but phantosmia and dysosmia are more common in trauma- than in NSD-related dysfunctions.

Although damage to higher neural pathways may contribute to these dysfunctions, it is believed that the most common mechanism is crushing or severing of the olfactory neuron axons at the point at which they pass through the small openings in the cribriform plate. Although histopathological studies in patients with posttraumatic anosmia have been limited, they seem to be consistent with traumatic severing of the olfactory filaments at the cribriform plate, followed by regeneration of the neuroepithelium and a failure of regenerating axons to reach the olfactory bulb (possibly as a result of fibrotic healing of the lamina of the cribriform plate and closure with scar tissue).

Aging, exposure to toxic compounds, and congenital deficiencies are other well-established etiologic factors in smell dysfunction. Together with the three major factors previously discussed, these probably encompass most instances of olfactory dysfunction in which a causal condition can be identified, although a number of medical conditions (most notably Parkinson's disease and Alzheimer's disease) and some medications have also been associated with smell disorders.

Given the relatively unprotected position of olfactory receptor neurons, it is not surprising they are susceptible to damage from pollutants in the ambient air. There are substantial animal toxicological data demonstrating damage to the olfactory neuroepithelium and bulb by airborne chemicals, as well as a large but scattered literature on the adverse effects of occupational exposures to industrial chemicals on the sense of smell. In humans, both acute and chronic exposures to a variety of chemical agents have been associated with olfactory dysfunction, which may be either temporary or permanent. This factor may play an important role in age-related smell loss.

There is still relatively little known about genetic/congenital olfactory dysfunction. Kallmann's syndrome is the principal genetic syndrome associated with anosmia, which is also characterized by hypogonadotropic hypogonadism (leading to a failure in sexual development). There are scattered reports of other forms of familial anosmia. However, frequently there is no family history of a similar problem, and it is impossible to rule out early childhood loss secondary to a head injury or upper respiratory infection in these cases.

Prognosis

Antibiotic, anti-inflammatory, and surgical therapy can reverse NSD-related smell loss. However, long-term management is complicated by the chronicity of the underlying disorders, and even patients who obtain effective treatment may be subject to recurring episodes of loss.

In cases of smell dysfunction secondary to peripheral nerve damage (such as is presumably the case in most instances of disorders that follow URI or toxin exposure, and in at least some cases of trauma-related disorder), spontaneous recovery is theoretically possible due to the regenerative capacity of the olfactory neural receptors. In fact, there is no question that some people do recover function, at least partially, following these kinds of insults, although in humans this typically appears to be a very gradual process, taking place over years. However, estimates of the incidence of full or partial recovery vary widely, from as low as 10% to as high as 67%.

No effective treatment for smell disorders has been identified, other than those associated with nasal/sinus disease. Zinc is often prescribed, but has been shown to be no more effective than a placebo in a double-blind study. A number of other treatments have been suggested, especially for problems secondary to URI, but controlled clinical trials have not been conducted.

Beverly J. Cowart

See also Ageing and Chemical Senses; Olfaction; Perceptual Development: Taste and Olfaction

Further Readings

Cain, W. S., Gent, J. F., Goodspeed, R. B., & Leonard, G. (1988). Evaluation of olfactory dysfunction in the Connecticut Chemosensory Clinical Research Center. *Laryngoscope, 98,* 83–88.

Cowart, B. J., Young, I. M., Feldman, R. S., & Lowry, L. D. (1997). Clinical disorders of smell and taste. *Occupational Medicine: State of the Art Reviews, 12,* 465–483.

Deems, D. A., Doty, R. L., Settle, R. G., Moore-Gillon, V., Shaman, P., Mester, A. F., et al. (1991). Smell and taste disorders, a study of 750 patients from the University of Pennsylvania Smell and Taste Center. *Archives of Otolaryngology, Head-Neck Surgery, 117,* 519–528.

Duncan, H. J., & Seiden, A. M. (1995). Long-term follow-up of olfactory loss secondary to head trauma and upper respiratory tract infection. *Archives of Otolaryngology, Head-Neck Surgery, 121,* 1183–1187.

Getchell, T. V., Doty, R. L., Bartoshuk, L. M., & Snow, J. B., Jr. (Eds.). (1991). *Smell and taste in health and disease.* New York: Raven Press.

Murphy, C., Schubert, C. R., Cruickshanks, K. J., Klein, B. E. K., Klein, R., & Nondahl, D. M. (2002). Prevalence of olfactory impairment in older adults. *Journal of the American Medical Association, 288,* 2307–2312.

Reden, J., Mueller, A., Meuller, C., Konstantinidis, I., Frasnelli, J., & Hummel, T. (2006). Recovery of olfactory function following closed head injury or infections of the upper respiratory tract. *Archives of Otolaryngology, Head-Neck Surgery, 132,* 265–269.

OLFACTION: EVOLUTION OF

Olfaction is generally defined as the sensory system that allows an organism's nervous system to detect chemicals coming from a distant source. Senses that require contact with the chemical source, such as when we taste food on our tongue, are not considered olfaction; neither is the kind of chemoreception present in individual cells or in single-celled animals. This definition is important to keep in mind when trying to understand the evolutionary history of the olfactory system. For example, lobsters can have chemosensory hairs on their antennae, antennules, legs, mouth parts, and carapace. To be able to make inferences about evolution, we need to know which of these groups of hairs constitutes the olfactory system. Thus, we need behavioral or physiological data demonstrating which hairs detect chemicals at a distance.

Such information is available for many vertebrates and for some invertebrate species. These data indicate that olfactory systems probably evolved many times independently, but that the olfactory systems in these distantly related animals have repeatedly evolved similar structures and strategies. Thus, the olfactory system in vertebrates, including humans, does not share deep evolutionary roots with olfactory systems in other groups of animals. Within vertebrates, large-scale changes have occurred in the genes that code for molecules that function as receptors for the chemicals we smell. A vomeronasal system emerged as a distinct olfactory subsystem in the ancestors of amphibians, reptiles, and mammals, and a system for regulating incoming olfactory information, the terminal nerve, emerged in early jawed vertebrates. Finally, reliance on olfactory information generally decreased in jawed vertebrates relative to our earliest ancestors, and again in Old World primates relative to other mammals. This entry will cover evolutionary comparisons across animal groups and evolutionary changes within vertebrates.

Evolutionary Comparisons Across Animal Groups

Animals are generally divided into 30 to 40 taxonomic groups, called *phyla* (singular: phylum). We only know about the organization and function of the olfactory system in four phyla: vertebrates; arthropods, especially selected species of insects, lobsters, and crabs; mollusks, largely limited to a few species of snails; and roundworms, in particular the nematode *Caenorhabditis elegans*. Although all these animals are capable of detecting chemicals from distant sources, their olfactory systems are probably not descended from one in a common ancestor. Instead, olfactory systems probably evolved repeatedly, and the similarities are due to the fact that the olfactory systems in these different animals have to perform a similar task, translating external odorant signals into internal neural signals.

Olfactory Sensory Cells

The appearance of the olfactory sensory cells in all these animals is strikingly similar. The sensory cells are neurons rather than the modified epithelial cells used in the visual, auditory, and taste systems. They have a single dendrite that is capped with minute hairlike extensions, called *cilia* or *microvilli,* that extend into a fluid-filled space. In insects, this fluid-filled space is under the cuticle, inside tiny hairs on the antennae. In terrestrial vertebrates, like humans, the fluid is produced by specialized mucus glands below the olfactory sensory epithelium. In some insects and vertebrates, this mucus contains specialized proteins called *odorant binding proteins.* The function of odorant binding proteins is unclear, but they may be involved in carrying odorants across the interface between air and mucus, or they could be involved in inactivating odorants after they have been detected. The odorant binding proteins in insects and mammals are completely different families of proteins, indicating that they evolved independently for use in olfaction.

Olfactory Receptors

The membranes of the cilia or microvilli contain molecules called *odorant receptors* that are activated when odorants bind to them. The molecular structure of the odorant receptors has been deciphered for nematodes, insects, and vertebrates, and they turn out to have interesting similarities and differences. All are large molecules that loop back and forth across the receptor membrane seven times. Because many receptors with this structure interact with a class of proteins inside the cell called *G proteins,* these receptors are generally called *G protein coupled receptors* (GPCRs). The family of GPCR molecules is huge and diverse, and is involved in many processes, including detecting light, neurotransmitters, hormones, and calcium. GPCRs may be particularly useful for detecting odorants because they activate biochemical cascades inside the cell that can greatly amplify the original signal, which would be ideal for detecting tiny numbers of odorant molecules. In addition, GPCRs possess a feature that is unusual among receptors, which is that a single receptor can activate different cascades when different molecules bind to its extracellular regions. This flexibility means that the olfactory receptor neurons in vertebrates and insects, which seem to express only one or a few odorant receptor genes, can still show different responses to stimulation by different odorants, allowing for detailed discrimination among stimuli. In contrast, the olfactory receptor neurons of nematodes seem to contain dozens of odorant receptor types, suggesting that the olfactory system of nematodes may be tuned to detect the presence or absence of certain classes of odorants, rather than making fine discriminations among odorants.

The odorant receptor genes show very low sequence similarity across phyla. In addition, recent studies indicate that the odorant receptors in insects may be “inside out” relative to the configuration seen in vertebrates and nematodes. Both of these findings indicate that GPCRs were probably co-opted independently to be used as odorant receptors in different phyla.

Beyond the Receptors

Beyond the sensory organ, the axons of olfactory receptor neurons in mollusks, arthropods, and vertebrates terminate in the brain in unusual tangles of fibers called *glomeruli.* Glomeruli probably evolved independently many times in olfactory systems of different animals, as they are present in vertebrates but not in our near relatives like lancelets and sea squirts. A careful analysis of the evolutionary relationships among arthropods suggests that glomeruli may have evolved many times in this phylum alone. Although the function of these structures is not clear, they seem to be important for processing olfactory information, as evidenced by the fact that they appear in different parts of the nervous system in different animals: For example, in some spiders that have olfactory hairs on their legs, glomeruli are found in the parts of the nervous system along the body of the animal that receive input from these hairs.

In vertebrates and insects, olfactory receptor neurons containing the same odorant receptor send their axons to the same glomeruli, indicating that these structures play a role in decoding odorant identity. The nature of this code is not known, however, and not all olfactory receptor neurons project to glomerular structures. Not only are glomeruli lacking in some arthropods, but even within an individual variability occurs. For example, in spiny lobsters (*Panulirus argus*), the olfactory neurons on one part of the antennule project to glomerular structures, but those on another part do not. The functional differences between these pathways are currently being explored, and such studies should shed light on the role of glomeruli in olfactory perception.

Evolutionary Changes Within Vertebrates

Changes Within Receptor Genes

In vertebrates, four families of genes have been shown to code for molecules that function as odorant receptors: the olfactory receptor (OR) genes, which are expressed in the olfactory epithelium in every vertebrate examined; the trace amino acid receptor (TAAR) genes, which were recently discovered to be expressed in the olfactory epithelium; and two groups of vomeronasal receptor genes, the V1Rs and V2Rs, which are expressed in the vomeronasal epithelium in frogs and mammals and in the olfactory epithelium in teleost fish.

The OR genes have been sequenced from a relatively large number of animals, allowing researchers to examine the evolutionary dynamics of the gene family. Some analyses indicate that the OR genes underwent dramatic changes when vertebrates crawled up on land, and that the gene family can be roughly divided into receptors used for smelling in water and others for smelling in air. If so, the genes for smelling in air were duplicated and diversified quickly, presumably allowing terrestrial vertebrates to discriminate among many different odorants. The size of the OR repertoire varies dramatically, even within tetrapods, the group of vertebrates that includes amphibians, reptiles, birds, and mammals. For example, the size of the OR repertoire is notably smaller in platypus than in other mammals, although it is not known whether this small repertoire was inherited from the earliest mammals or is the result of gene loss. Interestingly, the platypus also has the largest known repertoire of V1R genes, suggesting that, overall, their chemosensory abilities may not be much different than that of other mammals. The number of functional OR genes appears to be reduced in some primates, particularly in humans. The perceptual consequences are not clear, but the dramatic reduction in numbers of functional OR genes suggest that we may have reduced olfactory sensitivity or ability to discriminate among odorants, compared with other mammals and even other primates.

To date, the TAARs have been shown to be present in the olfactory epithelium only in mice, humans, and zebra fish. Although the OR repertoire can be extremely large, comprising hundreds or thousands of genes, only a handful of TAARs seem to be expressed in the olfactory epithelium. Other members of the family are used as neurotransmitter receptors in the brain. The discovery of TAAR expression in the nose raises the possibility that members of other families of neurotransmitter receptors may be expressed in the olfactory epithelium for use in odorant detection. Nevertheless, such genes are likely to be so few in number that they do not significantly compensate for large losses in numbers of OR genes, as is seen in humans.

The vomeronasal receptor genes, V1Rs and V2Rs, are not closely related to the ORs or to each other, and therefore were probably co-opted independently out of the larger GPCR family. Both families were first discovered in mice, and attempts to unravel their evolutionary history are just beginning. Nevertheless, V1Rs have now been described in many classes of vertebrates, including lampreys, bony fish, and amphibians, as well as mammals. In contrast, the V2Rs have not been found in lampreys, but have been described in bony fish, amphibians, and mammals; it is not clear whether the V2Rs evolved more recently than the V1Rs, or have changed so much that they are hard to find in lampreys. Curiously, the VRs are found in receptor neurons in the olfactory epithelium of fish, but are expressed in a separate sensory organ, the vomeronasal organ, in tetrapods.

Origin of the Vomeronasal System

In tetrapods, the olfactory and vomeronasal receptor neurons send their axons to separate regions of the forebrain, the main and accessory olfactory bulbs. It is not yet known how the olfactory and vomeronasal systems, which seem to be mixed together in fish, became separated in tetrapods. The separation of these two systems may have led to increased sensitivity to particular odorants, or greater ability to discriminate among certain types of odorants. It is difficult to speculate about the perceptual consequences of this separation because researchers do not understand the function of the vomeronasal system. In addition, the vomeronasal system has been lost many times independently, notably in alligators and crocodiles, birds, whales and dolphins, some salamanders, several groups of bats, apes, and humans. Although the VRs have not yet been examined in all these groups of animals, data from the human genome indicate that all the VRs in humans are pseudogenes and therefore not functional.

Origin of the Terminal Nerve

Jawed vertebrates possess a terminal nerve, which extends between the forebrain and olfactory organs. The terminal nerve appears to tune the olfactory receptor neurons to produce optimal responses given the animal's behavioral or physiological state. For example, when animals are hungry, a chemical released by the terminal nerve causes the olfactory organ to respond more strongly to food odorants than it does when the animal is full. The terminal nerve also releases chemicals that change responses of olfactory receptor neurons depending on the animal's breeding condition. In bony fish, an additional branch of the terminal nerve extends to the retina, and incoming odorant information can alter visual sensitivity. This type of filtering, which occurs right at the level of the sense organ, can play a strong role in shaping perception by allowing only certain types of olfactory information to be carried to the brain or by changing the relative significance of particular odorants. It is not yet clear whether jawless fish (hagfish and lampreys) lack the type of filtering provided by the terminal nerve or use different means of filtering incoming olfactory information.

Proportion of the Brain Devoted to Olfaction

One measure of the relative importance of a sensory modality is the proportion of the brain dedicated to processing that type of sensory information. Thus, researchers have tried to determine the extent of olfactory projections in the forebrain, which contains the cortex in mammals and its equivalent in other groups of vertebrates. Although the proportion of the forebrain devoted to processing olfactory information seems to be fairly constant in jawed vertebrates (sharks and their kin, bony fish, and tetrapods), olfactory projections appear to be more extensive in jawless fish—hagfish and lampreys. Because the first vertebrates were also jawless, this observation suggests that early vertebrates relied heavily on olfaction as they navigated, foraged or hunted, and sought mates. Compared with our ancestors, the relative importance of olfactory information is probably somewhat reduced in sharks and their kin, bony fish, and tetrapods, including humans.

Because brain capacity is limited, emphasis on one sensory system may necessitate reduction in the importance of others. Thus, many primates, and particularly humans, may have a poor sense of smell because we rely heavily on our visual system. Although this phenomenon has been hard to document at the anatomic level, molecular data lend support to this idea. Specifically, trichromatic (three-color) vision has evolved independently in Old World primates, such as macaques, baboons, and apes, and in one group of New World monkeys, the howler monkeys. Interestingly, in both groups, the number of olfactory receptor genes that code for working receptors is significantly decreased, suggesting that improved vision may indeed reduce reliance on olfactory information.

Heather L. Eisthen

See also Auditory System: Evolution of; Evolutionary Approach: Perceptual Adaptations; Eyes: Evolution of; Olfaction; Olfactory Receptors and Transduction; Perceptual Development: Taste and Olfaction; Visual System: Evolution of; Vomeronasal System

Further Readings

Ache, B. W., & Young, J. M. (2005). Olfaction: Diverse species, conserved principles. *Neuron, 48,* 417–430.

Eisthen, H. L. (2002). Why are olfactory systems of different animals so similar? *Brain, Behavior and Evolution, 59,* 273–293.

Halpern, M., & Martinez-Marcos, A. (2003). Structure and function of the vomeronasal system: An update. *Progress in Neurobiology, 70,* 245–318.

Kaas, J. H. (Ed.). (2006). *Evolution of nervous systems.* London: Elsevier Press.

Kambere, M. B., & Lane, R. P. (2007). Co-regulation of a large and rapidly evolving repertoire of odorant receptor genes. *BMC Neuroscience, 8*(Suppl. 3), S2.

Strausfeld, N. J., & Hildebrand, J. G. (1999). Olfactory systems: Common design, uncommon origins? *Current Opinion in Neurobiology, 9,* 634–639.

OLFACTION: FEATURE DETECTION AND INTEGRATION

In our everyday lives we experience a large variety of smells, each one comprising several to thousands of different molecules. Our olfactory system has the amazing ability to detect and discriminate these odors based on the molecules that they comprise. Odor processing and recognition takes place as odor information passes through several levels of the olfactory system. Beginning in the nose, odors are broken up into a collection of individual parts, or features. This feature information is refined in the olfactory bulb and is then synthesized into a single odor percept within the cortex. In this entry, odorant feature detection and the synthesis of features in odor objects will be covered.

Odorant Feature Detection

Every time we inhale a breath, volatile molecules emitted from odor sources travel up our nostrils and interact with our olfactory epithelium. Located within the epithelium is a thin sheet containing millions of olfactory sensory neurons. Each one of these neurons expresses a single receptor protein type that recognizes and binds with certain molecular structures or features of odorant molecules, such as hydrocarbon chain length or functional group. This feature allows individual odor molecules to interact with a large number of different receptor types in the epithelium. The unique set of receptor neurons that are activated by an odor forms the basis of an odor identity code that is projected to the olfactory bulb.

Within the bulb, sensory neuron axons synapse with bulbar output neurons in distinct spherical regions known as *glomeruli* that spread out across the entire surface of the olfactory bulb. Each glomerulus only receives input from sensory neurons expressing the same receptor type. This organization allows for odors to be represented in the initial stages of olfactory processing as a set of glomeruli based upon the set of receptors activated by that odor. Thus, each odor creates a specific topographical map of glomerular activity across the surface of the olfactory bulb. However, these maps can be become quite complex and highly overlapping, especially when odors contain similar molecules or features. In this case, the input map of glomerular activity must be fine-tuned by subsequent processing in order for similar odors to be perceived as distinct.

This process begins at the first synapse in the olfactory bulb with a complex network of excitatory and inhibitory neurons that surround each glomerulus (see color insert, Figure 9). This interneuronal network affects odor maps in two main ways. First, a subset of periglomerular neurons that receive direct sensory input provide feedback inhibition back onto the sensory axons within each glomerulus. This is accomplished by blocking neurotransmitter release from the activated sensory neurons and effectively damping the input. This mechanism serves as a way for the system to control the overall strength of the initial input into the bulb. Second, another group of juxtaglomerular neurons excite the inhibitory neurons surrounding adjacent glomeruli, forming a lateral inhibitory network. This center-surround inhibition allows strongly activated glomeruli to block or reduce the excitation of the output neurons associated with other similarly responding glomeruli. By only allowing the strongly activated sensory neurons to pass on their odor information to higher processing levels, glomerular inhibition serves to reduce representational overlap between similar odors and sharpen odor maps at the glomerular level.

Beyond the glomerular layer lie the main output neurons of the bulb, mitral cells and tufted cells. Each of these cells projects a primary dendrite into a single glomerulus, where it synapses with sensory axon terminals. Thus, all the mitral and tufted cells associated with an individual glomerulus receive exclusive input from a single receptor type. It is at this point that the spatial odor map at the glomerular layer is transformed into a pattern of spiking output at the mitral/tufted cell level. Similar to sensory neurons, mitral and tufted cells respond to receptor input with trains of action potentials that are, in many cases, highly correlated with respiratory cycles. Mitral and tufted cells have overlapping but individually distinct odorant receptive fields. For example, recordings from mitral/tufted cells in anesthetized rodents stimulated by a wide variety of odorants have shown that individual cells tend to respond to a range of odor molecules that share similar structural features, including carbon chain length and type of functional group. Although mitral/tufted cells function as molecular feature detectors that reflect the input of the sensory neurons to which they synapse, their odor responses are further processed through an extensive network of inhibitory granule cells.

In addition to a primary dendrite, mitral/tufted cells project several lateral dendrites that extend radially below the glomerular layer. Along the way, they synapse with the dendrites of many granule cells. The lateral dendrite projections are extensive, with some cells projecting dendrites nearly halfway across the bulb. These long projections allow mitral/tufted cells to interact with a large population of granule cells both locally and more globally. The synapses between mitral/tufted and granule cell dendrites are unique in that they are reciprocal. In this case, mitral/tufted cells release the excitatory neurotransmitter glutamate onto granule cells and cause them to release the gamma-aminobutyric acid (GABA) inhibitory neurotransmitter back onto the mitral/tufted cells.

These synapses serve several functions, the first of which is to gate mitral/tufted cell activity through feedback inhibition. In this way, excited mitral/tufted cells fire action potentials that travel down their lateral dendrites and excite a population of granule cells. GABA release from these granule cells activates inhibitory currents in the mitral/tuft cell soma that would stop or reduce mitral/tufted cell firing. Although the exact function of this mechanism is not known, it could serve as a control mechanism similar to inhibition at the glomerular level.

Another function of this synapse is to allow for groups of mitral/tufted cells to inhibit each other through lateral inhibition. Mitral/tufted cell lateral dendrites can extend quite far across the bulb, making contacts with many granule cells along the way. Some of theses granule cells synapse with mitral/tufted cells associated with other glomeruli with similar receptive fields. In this case, mitral/tufted cells that are strongly activated by an odor would fire action potentials that travel down their lateral dendrites and excite a population of granule cells. The granule cells would then, in turn, inhibit other sets of mitral/tufted cells that are only weakly activated by the odor. This mechanism could serve to enhance the contrast of strongly activated output neuron responses and ultimately sharpen mitral/tufted cell odor receptive fields. There is evidence of this in electrophysiological recordings from mitral/tufted cells. These experiments have demonstrated that individual cells are inhibited by groups of odors that are structurally similar to the odors that excited the cells. Furthermore, this inhibition could be blocked by the application of drugs that block granule cell glutamate receptors.

Granule cell inhibition can also serve to alter the timing of mitral/tufted cell odor responses. Through mechanisms similar to those previously mentioned, populations of excited granule cells could provide global, simultaneous inhibition of mitral cell ensembles. This would serve to synchronize the spiking output of mitral/tufted cells activated by the odorant. This synchronized excitation–inhibition can be observed as field potential oscillations in the presence of odor input. The synchronized firing of all mitral/tufted cells activated by an odor could serve to temporally bind all the feature information extracted at the receptor level into a specific spatiotemporal pattern of output activity that can be interpreted by the olfactory cortex.

The shaping of mitral/tufted cell responses by inhibitory networks within the bulb can be heavily influenced by neuromodulatory input from other brain regions. In general, most of these inputs are focused onto the periglomerular and granule cells and have direct effects on the level of mitral/tufted cell inhibition. For example, at the glomerular level, serotonin release from fibers originating in the brain stem enhances periglomerular cell mediated presynaptic inhibition. While at the granule cell level, cholinergic input from fibers originating in the basal forebrain can increase granule cell excitability and therefore enhance their inhibition onto mitral/tufted cells. Changes in bulbar norepinephrine levels can also alter mitral/tufted cell firing rates, presumably by altering granule cell properties. Finally, granule cells also receive centrifugal input from the olfactory cortex. These inputs are excitatory and presumably allow for cortical control of bulb inhibitory levels.

Synthesis of Features in Odor Objects

Despite the fact that everyday odors consist of many different molecules and molecular features, we tend to experience them as whole perceptual objects and not as a collection of individual features or components. This process occurs through the transformation of highly analytical afferent input into a synthetic whole-odor representation in the cortex. This synthesis is made possible by to the unique architecture and learning capabilities of the olfactory cortex.

The olfactory cortex is defined as those areas receiving direct synaptic input from the olfactory bulb. This includes several separate cortical areas, the largest of which is the piriform cortex. The primary afferent input into the piriform cortex comes from mitral/tufted cells by way of the lateral olfactory tract. Here, mitral/tufted cell axons spread out across the entire cortical surface and form excitatory synapses onto the primary dendrites of pyramidal cells, the principle cortical neuron type. Although there is some topographical specificity to these mitral/tufted cell inputs, no direct point-to-point transfer of the spatial glomerular map onto the cortex has been observed. The broad, overlapping distribution of mitral/tufted cell input onto the piriform cortex allows each pyramidal cell to receive converging input from several mitral/tufted cells that are associated with different glomeruli. Immediately deep to the mitral/tufted cell input layer lies a layer of intracortical association fibers. In this layer, pyramidal cell axon collaterals branch out extensively and synapse with the primary dendrites of thousands of other pyramidal cells. Similar to mitral/tufted cells, this input is excitatory and capable of driving pyramidal cell activity.

This anatomy allows the piriform cortex neurons to function not as feature detectors like in the bulb, but as feature integrators that are excited by specific combinations of different odor features. Several lines of evidence support this idea. For example, activity in the cortex is much more widely distributed, with odors evoking large, overlapping patterns that do not appear to reflect odor identity. Pyramidal cells have lower firing rates and display less odor-evoked inhibition. Individual cell receptive fields tend to be broader than those of mitral/tufted cells, and in some cases, include odors belonging to different structural categories. Finally, unlike mitral/tufted cells, pyramidal cells display rapid, odor-specific adaptation.

The synthetic nature of odor coding in the piriform cortex is heavily dependent on prior experience and learning. The combination of diffuse afferent input and extensive intracortical connectivity has been hypothesized to function as an autoassociative network, whereby patterns of input are stored and retrieved through alterations in the strength of association fiber connections. In this case, the synaptic strength of association fiber input onto active pyramidal cells increases, allowing for long-lasting increases in the ability of the association fibers to subsequently activate the cell. Strengthening the connections between specific ensembles of pyramidal cells encoding an odor allows for a distributed network of cells to rapidly and reliably identify odors. In addition, stable representations of odors could easily be recalled within the cortex, even if the input was degraded by changes in concentration or contaminated by other background odors.

Finally, the piriform cortex receives input from neuromodulatory regions that can directly affect cortical odor representation. For example, cholinergic fibers from the basal forebrain project into the cortex and terminate in the pyramidal cell layer. At the cellular level, acetylcholine release from these fibers has major effects on pyramidal cell activity, including increasing cell excitability and modulating association fiber long-term plasticity. These effects enhance the plasticity of the piriform cortex synapses and play a major role in odor recognition, memory, and discrimination.

Max. L. Fletcher

See also Olfaction; Olfactory Central Processing; Olfactory Receptors and Transduction; Perceptual Development: Taste and Olfaction

Further Readings

Aungst, J. L., Heyward, P. M., Puche, A. C., Karnup, S. V., Hayar, A., Szabo, G., et al. (2003). Centre-surround inhibition among olfactory bulb glomeruli. *Nature, 426,* 623–629.

Haberly, L. B. (2001). Parallel-distributed processing in olfactory cortex: New insights from morphological and physiological analysis of neuronal circuitry. *Chemical Senses, 26,* 551–576.

Kadohisa, M., Fletcher, M. L., & Wilson, D. A. (2006). Olfactory cortex. *Seminars in Cell and Developmental Biology, 17,* 462–470.

Uchida, N., Takahashi, Y. K., Tanifuji, M., & Mori, K. (2000). Odor maps in the mammalian olfactory bulb: Domain organization and odorant structural features. *Nature Neuroscience, 3,* 1035–1043.

Wilson, D. A., & Stevenson, R. J. (2006). *Learning to smell: Olfactory perception from neurobiology to behavior.* Baltimore: Johns Hopkins University Press.

Yokoi, M., Mori, K., & Nakanishi, S. (1995). Refinement of odor molecule tuning by dendrodendritic synaptic inhibition in the olfactory bulb. *Proceedings of the National Academy of Sciences USA, 92,* 3371–3375.

OLFACTION AND REPRODUCTIVE BEHAVIOR

Sexuality and successful reproduction are orchestrated by three kinds of olfactory signals, which operate either with or without our awareness: odors, pheromones, and vasanas. These three types of olfactory signals will be discussed in this entry. Odors are, by definition, molecules at relatively high concentrations that are detected consciously as smells. With experience, odors can become associated with sexual responses (conditioned stimuli), different molecules with a similar smell elicit similar responses (generalization), and yet, as time goes by, the molecules are no longer detectable (habituation). Pheromones, in contrast, operate in minute amounts, without needing to be consciously detected as an odor, and change behavior or hormone function. They are species-specific signals without generalization or habituation and don't become conditioned stimuli. Intermediate are vasanas, a newly recognized functional class of molecules that also affect mood, attention, and physiology in small amounts without conscious detection, but nonetheless have some odorlike properties.

Odors

Sexuality

Body Scents

People have mild body scents that are different in men, women, children, teenagers, and grandparents, and individuals also have their own odor signature. Not only can animal species with an exquisite sense of smell (such as dogs and rats) readily distinguish these different odors, so can people. Indeed, for women, the quality of a person's odor is often the most important attribute in a sexual relationship, whereas appearance is most important for men.

The perception of body scents and odors is highly culture bound. In America, strong body odors produced, for example, during exercise are viewed as unpleasant, driving a multibillion-dollar industry to remove or mask them; even mild body scents are taboo. Yet, in the arts and in other cultures, body scents, if not strong body odors, are not only considered natural but celebrated.

In humans, one source of a person's odor signature is the proteins expressed on cell surfaces, termed the human leukocyte antigen, or generally for most species, the major histocompatibility complex (MHC). The MHC has been studied primarily in the context of its immune function, enabling a cell to be recognized as "self" rather than an invading pathogen. By smell alone, humans can detect differences in genes for the MHC. Within inbred groups, and when sexuality is made explicit, people prefer scents from other people with fewer matches to their own MHC genes. In outbred populations, where people typically have no matches, people avoid scents from people with zero matches, preferring those from people with at least a few MHC matches. Interestingly, the odors are not more or less familiar, but instead pleasantness drives the preference. This suggests that processing of MHC information involves olfactory projections to the amygdala, which responds differentially to the emotional value of an odor rather than more circuitous odor recognition, verbal identification, and cognitive processing.

Odor Conditioning

Through conditioning and associative learning, both natural body odors and serendipitous environmental odors play an integral role in sexual behavior of most mammals. This principle has been elegantly demonstrated in rodents. For example, male rats respond to the odor of a female in heat. If he has mated with females anointed with lemon oil, then lemon oil cues sexual arousal. Likewise, when a male fetus in the womb is exposed to lemon oil, then in adulthood he prefers females scented with lemon oil. Conversely, females prefer odors of sexually mature males, and can also learn arbitrary odors as sexual cues. If an odor is associated with negative stressful experiences, then it alone can impair male and female sexual function.

Similarly, in humans who have suffered sexual abuse or a rape, odors that were present during the assaults will often trigger a panicked response, becoming part of a posttraumatic stress disorder. Despite copious literary allusions, however, there is little empirical research to date demonstrating a positive role for odor-associations in human sexuality.

Parenting

Odors also play a strong role in parenting, the second primary component of reproduction. In some species, such as house mice, parents recognize their genetic offspring by odor alone, even if they were raised apart until adulthood. In other species, such as pigs, newborns are readily adopted immediately after birth, and close interactions make their odors as easily recognizable as the parent's own genetic offspring. Human mothers need to smell their newborn babies for only an hour or less to enable them to pick out the odor of their own baby from odors of other newborn infants. The attractiveness of their baby's odors, however, is enhanced by previous maternal experience and high postpartum levels of the stress hormone cortisol.

Pheromones

From Insects to Humans

Pheromones are social chemosignals operating in minute amounts, typically many orders of magnitude less than is required to produce the perception of an odor. They enable a member of a social group to change the nervous system or physiology of another group member, regulating specific aspects of its behavior, neuroendocrine function, or development.

Pheromones were first defined in insects in the late 1950s, after a specific molecule, called bombykol, was identified that served as a sex attractant for the silkworm moth. A female moth releases her pheromones in a plume, and just a few molecules will trigger neural programs in the male, causing him to navigate upwind for miles to find and mate with her. In the past half-century, the range of species utilizing pheromones now extends from yeast to humans, and their reproductive functions extend from sexual maturation to modulation of mood and attention. The description in the following section includes three types of pheromones used in reproduction.

Releaser, Primer, and Modulator Pheromones

When a boar sprays a sow with his saliva carrying releaser pheromones, he immediately triggers her lordotic mating stance, allowing him to mount. In general, releaser pheromones operate within seconds to minutes, triggering neural systems and relatively stereotyped behavior. Despite marketing claims, there is no evidence that humans have releaser pheromones; human sexuality is multidetermined and fundamentally context dependent.

Humans do utilize primer pheromones, which act over days or weeks to change neuroendocrine function. Women produce ovarian pheromones that regulate the timing of ovulation in other women. Specifically, they release different ovarian pheromones in their axillae during different phases of the menstrual cycle, which either enhances or delays the preovulatory surge of luteinizing hormone (LH) by accelerating or slowing LH pulses in recipient women. These ovarian pheromones likely mediate the menstrual synchrony that can develop among groups of women living together under specific conditions. Axillary compounds from men can also accelerate LH pulses, which may explain why women who rarely see men are more likely to have irregular menstrual cycles.

Primer pheromones are also produced during pregnancy and lactation, which regulate the fertility of other females. Indeed, breastfeeding pheromones from lactating women and their infants regulate follicular development during the ovarian cycles of women exposed only to breastfeeding compounds from mothers and infants.

There are many other aspects of fertility and reproductive behavior that are regulated throughout the reproductive lifespan. In rodents, puberty is regulated by pheromones from adult males and females in the social group. Pregnancy can be blocked by pheromones from a strange male. Sperm count is increased by the presence of a fertile female. And reproductive senescence can be delayed by pheromones from a young adult partner.

Modulator pheromones affect essential aspects of reproductive behavior, but do not act alone. Modulator pheromones from a male rat increase the intensity of the female's arched-back lordosis reflex after it has been elicited by his mount and flank palpation. Unlike releaser pheromones in male domestic pigs, however, a modulator pheromone is not sufficient to trigger the female rats' mating posture. Thus, modulator pheromones adjust ongoing behavior that is already in progress in response to a rich set of other physical and social stimuli.

In humans, modulator pheromones may enhance sexual motivation and regulate how someone interprets a situation, but without trigger specific thoughts or actions. Modulator pheromones from lactating women and their nursing infants increase sexual desire in fertile women, particularly during natural ebb. It does not increase sexual behavior, which is determined by their life circumstances.

Vasanas

Vasanas are an intriguing new class of olfactory compounds, which were first identified in the context of human sexual and social interactions, but may turn out to regulate other aspects of behavior and include a wide variety of environmental compounds. Androstadienone (4, 16-androstadien-3-one) is a steroid in human sweat, originally claimed to be a human releaser pheromone, triggering a feeling of relaxation during a romantic interaction with a man.

Androstadienone in fact affects more fundamental psychological states of men as well as women, such as attention to emotions. In the laboratory, it improves women's attention and mood; but for men only their attention is increased, while their mood gets worse. Thus, the outcome of increased attention depends on how a person is already reacting to their social situation. Only nanomolar amounts of androstadienone are enough to produce these particular effects, well below the threshold for odor detection.

Nonetheless, at higher concentrations, androstadienone does have an odor and a human olfactory receptor has been identified. As an odor, androstadienone increases sexual arousal when men and women view erotic films; this effect could result from odor conditioning, and contribute to its other effects when it is undetectable as an odor. This demonstrates how the same molecule can have different functional effects, depending on context and concentration. Many other molecules, both from the body and the environment, may also function both at the level of odors and subliminally as vasanas.

Martha K. McClintock

See also Attractiveness; Consciousness; Context Effects in Perception; Olfaction; Pheromones; Unconscious Processes

Further Readings

Jacob, S., McClintock, M. K., Zelano, B., & Ober, C. (2002). Paternally inherited HLA alleles are associated with women's choice of male odor. *Nature Genetics, 30,* 175–179.

Keverne, E. B. (2004). Importance of olfactory and vomernasal systems for male sexual function. *Physiology & Behavior, 83*(2), 177–187.

Koyama, S. (2004). Primer effects by conspecific odors in house mice: A new perspective in the study of primer effects on reproductive activities. *Hormones and Behavior, 3,* 303–310.

McClintock, M. K. (1998). Whither menstrual synchrony? *Annual Review of Sexual Research, 9,* 77–95.

McClintock, M. K. (2002). Pheromones, odors and vasanas: The neuroendocrinology of social chemosignals in human and animals. In A. Arnold, A. Etgen, D. Pfaff, S. Fahrbach, & R. Rubin (Eds.), *Hormones, brain and behavior* (Vol. 1, pp. 797–870). San Diego, CA: Academic Press.

Restrepo, D., Arellano, J., Oliva, A. M., Schaefer, M. L., & Lin, W. (2004). Emerging views on the distinct but related roles of the main and accessory olfactory systems in responsiveness to chemosensory signals in mice. *Hormones and Behavior, 3,* 247–256.

Snowdon, C. T., Ziegler, T. E., Schultz-Darken, N. J., & Ferris, C. F. (2006). Social odours, sexual arousal and pairbonding in primates. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences, 361,* 2079–2089.

Vosshall, L. B. (2008). Scent of a fly. *Neuron, 59,* 685–689.

Yamazaki, K., & Beauchamp, G. K. (2007). Genetic basis for MHC-dependent mate choice. *Advances in Genetics, 59,* 129–145.

OLFACTOMETRY

Olfactometry is a somewhat misleading term. Although it implies that it relates to the measurement of olfactory sensitivity, it is used to describe a technique that allows presentation of odorous stimuli so that the subjects' olfactory function can be assessed. Thus, olfactometry relates to methods of stimulus presentation. Typical problems described in this entry that are associated with odor presentation include the control of concentrations reaching the olfactory receptors in the olfactory cleft, the change of stimulus concentration over the time of administration time, or the stability of odor concentrations during repeated exposures.

For many psychophysical studies, odors are presented, for example, in glass jars, in odorized pens, or as microencapsulated odors glued to paper that are released through scratching them with a pencil. There is detailed research on the influence of numerous variables on the results obtained with the various techniques (e.g., effects of environmental temperature or humidity, size and material of jars, consequences of solvent on the release of odors). These relatively simple techniques of odor presentation have been successfully employed in most studies related to human olfactory function. Drawbacks of these techniques include the relatively poor control of delivered stimulus concentrations, or the need for the subjects' or patients' cooperation.

For some questions, however, more sophisticated techniques are desirable. For example, to know how much odor is entering the nose at what time and for how long most researchers would employ an *olfactometer,* a stimulator releasing odors in a controlled fashion. A number of such olfactometers have been constructed.

The following describes an olfactometer that works on the principles of air-dilution olfactometry. Developed by Gerd Kobal, this olfactometer produces odorous stimuli with rapid onset; stimuli are controlled in terms of timing, duration, and intensity; and the presentation of stimuli does not simultaneously activate sensory systems other than the olfactory system because odors are applied intranasally by means of a cannula (inner diameter 4 millimeters, mm)—typically independent of the respiratory cycle. This cannula has to be inserted into the nostril so that its opening lies beyond the nasal valve. Odor pulses are embedded in a constantly flowing airstream. This constant airstream has a typical flow rate ranging 2 to 8 liters per minute (1/min). The higher the flow rate, the better the stimulus presentation has to be controlled.

Monomodal chemosensory stimulation (meaning specific stimulation of the olfactory or trigeminal nerves without concomitant stimulation of nasal mechanosensors) is achieved by directing two airstreams toward the olfactometer's outlet. Both airstreams have the same flow rate, same temperature, and same humidity. One of the two airstreams contains odor in a certain concentration; the other airstream is odorless. At the stimulator's outlet, a precisely tuned vacuum is attached so that the odor-containing airstream is vacuumed away and only odorless air is presented to the subjects. For stimulation, the vacuum is switched so that the odorless air is now sucked away and the odorized air reaches the subjects' nose. This mechanism allows the researcher to switch between odorless and odorized air within a very short time (2/3 of maximum concentration is reached within less than 20 milliseconds, ms). Different odor concentrations are generated with air dilution so that a pre-established, fully odor-saturated airstream is mixed into an odorless airstream.

The constant airflow directed into the subjects' noses requires humidification (≥80% relative humidity) and thermo stabilization (36 °C) because dry, cool air produces nasal congestion, mucus discharge, and pain, which interferes with smelling. Cross-contamination of odors needs to be prevented in olfactometers, otherwise all stimuli would smell similar after a while. In the Kobal-type olfactometer, a system of finely tuned airstreams is applied so that a small current of odorless air is preventing molecules from odorized tubing to be drawn to other tubing of the system. This also refers to the problem of background odor, which, when present during the interstimulus interval, will have a major effect on the outcome of measurements. The same applies to psychophysical testing with less sophisticated equipment—the presence of background or environmental odor will significantly affect olfactory measurements.

In commercially available olfactometers, valves and mass-flow controllers are typically under computer control. This allows the setup of sequences of stimuli with different quality, intensity, duration, or interstimulus interval, which helps to standardize even complex measurements—which is of utmost importance in a clinical setting.

What happens when olfactory stimuli are puffed into the nasal cavity? The result is a mixed activation of both nasal mechanosensory receptors and the olfactory system, which leads to numerous interactions. These interactions between the nasal somatosensory/trigeminal and olfactory system are difficult to predict.

Important questions within the context of olfactometry also relate to the choice of odor (whether it activates the trigeminal system or not, how familiar subjects are with the stimulus, whether it is associated with foods or not, etc.), and to stimulus duration and interstimulus intervals. Further, it is a matter of debate whether stimuli should be presented in synchrony with inspiration. Although there is no single answer to these questions, these concerns need to be carefully considered individually for each research project.

Thomas Hummel

See also Olfaction; Olfactory Stimulus

Further Readings

Doty, R. L., & Laing, D. G. (2003). Psychophysical measurement of human olfactory function, including odorant mixture assessment. In R. L. Doty, (Ed.), *Handbook of olfaction and gustation* (2nd ed., pp. 203–228). New York: Marcel Dekker.

Hummel, T., & Livermore, A. (2002). Intranasal chemosensory function of the trigeminal nerve and aspects of its relation to olfaction. *International Archives of Occupational and Environmental Health, 75,* 305–313.

Johnson, B. N., & Sobel, N. (2007). Methods for building an olfactometer with known concentration outcomes. *Journal of Neuroscience Methods, 160,* 231–245.

Kobal, G. (1981) *Electrophysiological investigations of the human sense of smell* [*Elektrophysiologische Untersuchungen des menschlichen Geruchssinns*]. Stuttgart, Germany: Thieme Verlag.

Lorig, T. S., Elmes, D. G., Zald, D. H., & Pardo, J. V. (1999). A computer-controlled olfactometer for fMRI and electrophysiological studies of olfaction. *Behavior Research Methods Instruments & Computers, 31,* 370–375.

OLFACTORY ADAPTATION

Walking into a barnyard can be an overpowering experience. The rich and varied smells one experiences are the result of molecules being inhaled up the nose and activating our olfactory system—the sense of smell. Fortunately, and perhaps surprisingly, this barnyard bouquet of smells fairly rapidly vanishes. After a few minutes of exposure, we no longer smell those odors that were so overwhelming on our arrival. The odor molecules are still present, but our response has diminished. This is called *olfactory adaptation.*

All sensory systems are able to adjust their responsiveness to input based on recent stimulation history and context. In olfaction, repeated or prolonged odor stimulation can lead to adaptation and a decrease in how strongly the sensory system responds. This decrease in the brain's response to odors causes a decrease in our perception of how strong the odor is, and, in some cases, can lead to us not smelling the odor at all. The magnitude of olfactory adaptation depends on the parameters of stimulus intensity, duration or rate of exposure, the identity of the specific stimulus involved, and, in many cases, the context of the exposure. The perceptual changes that occur during adaptation are caused by changes both in the nose and within the central nervous system circuits. Because olfactory adaptation can be very odor specific, it not only allows filtering of stable, noncritical odors (like the smelly barnyard), but also contributes to identification of new odors presented against that odorous background (such as a barnyard rose). This entry discusses the behavioral characteristics, neurobiology, and consequences of olfactory adaptation.

Behavioral Characteristics

Historically, there have been two terms to describe a decrease in response to a stimulus following repeated presentation of that stimulus: *adaptation* and *habituation.* The original use of these terms defined the decrease in behavioral response as habituation, and one potential neural mechanism of this habituation—a decrease in the response of those neurons that initially received the stimulus (e.g., receptor neurons in the eye, skin, or nose) as adaptation. However, it has become increasingly clear that behavioral habituation may involve adaptation at both the receptors and within the central nervous system. This has led to a blurring of the lines between what is meant by habituation and adaptation. Within the sensory physiology and psychophysics literature, *adaptation* has become the more common term to describe both decreased behavioral response and its potential underlying neural cause. Within the learning and memory literature, habituation and adaptation still tend to be used in somewhat different contexts closer to their historical usage. Here, olfactory adaptation refers to both the reduced behavioral/perceptual response and its presumed underlying neural cause.

Either repeated or prolonged exposure to an odor reduces subsequent responsiveness to that odor. This adaptation can be measured as either an increase in detection threshold (i.e., how strong the odor needs to be before it is noticeable), or as a decrease in response magnitude to strong odors (i.e., the intensity rating of an odor that is detectable). Depending on the conditions, as described next, olfactory adaptation can occur within seconds or minutes of exposure, or may take much longer if stimulus presentations are spaced over time or depending on the context or odor identity.

Sensory adaptation has several defining features beyond a response decrease. First, under most circumstances, sensory adaptation spontaneously recovers with time, though depending on the stimulation conditions, the duration of adaptation may vary from seconds to hours or days. Depending on the stimulus, detectable olfactory adaptation in both invertebrates and vertebrates can occur with less than one minute of exposure, and adaptation increases in magnitude as stimulus duration is prolonged. Spontaneous recovery from adaptation similarly can occur within seconds, though with longer exposure, full recovery may take weeks. For example, several studies have examined how people respond to daily exposure to odors in the workplace, like industrial solvents, plastics, and others. Daily exposure to these odors for six to eight hours a day for three weeks caused adaptation to the odors that required two weeks after removal of the odor to reverse.

Second, the rate and/or magnitude of adaptation is dependent on stimulus intensity. Adaptation rate and magnitude can be difficult parameters to isolate from each other, and relatively little data on olfactory adaptation rate (amount of change/time) exist. However, as odor concentration increases, the rate and/or magnitude of adaptation decreases. Thus, after a given duration of exposure, a strong stimulus will have adapted less than a weak stimulus (though see further examples). For example, it will take someone a long time to adapt to the odor from a bottle of lemon flavoring. The smell of lemon will be clear even after repeated sniffs from the bottle. However, we will fairly rapidly adapt to the gentle citrus scent coming from a lemon tree or a room freshener.

A third characteristic of olfactory adaptation is stimulus specificity. The degree of cross-adaptation from one odor to another has been used as a measure of structural or perceptual similarity between stimuli. In general, molecularly distinct odors (e.g., odors resulting from smelling molecules that are very distinct in shape or size) induce little cross-adaptation, meaning adaptation to one stimulus leaves responses to the other intact, while molecularly (or in some cases perceptually) similar odors induce greater cross-adaptation. Interestingly, the amount of cross-adaptation between two odors is often not reciprocal. That is, adaptation to odor A may induce strong adaptation to odor B, but adaptation to B may induce only minimal cross-adaptation to A. For example, adaptation to pentanol, a sweet, alcohol-like odor, causes strong cross-adaptation to propanol, another alcohol-like odor. In the converse, however, adaptation to propanol produces much less cross-adaptation to pentanol. Specific mechanisms contributing to these nonreciprocal interactions are unknown.

The degree of cross-adaptation can also be influenced by familiarity through perceptual learning. That is, familiar odors are more easily discriminated and show less cross-adaptation than novel odors. For example, rats show strong cross-adaptation between ethyl esters varying by one hydrocarbon, such as ethyl butyrate and ethyl valerate, if they have never been exposed to those odorants before. However, if they have been given prior exposure to those odorants, cross-adaptation is significantly reduced.

Fourth, the degree of adaptation to odors is influenced by nonolfactory factors, such as expectations, hunger, or stimulation from other sensory systems. For example, subjects told that a chemical with a particular odor may be toxic with prolonged exposure adapt significantly less to that odor compared to subjects told that the same odor is used with beneficial effects in aromatherapy. Thus, the context in which the odor was presented, and cognitive factors such as attention and expectations of the subject affect the rate or magnitude of odor adaptation. Similar results have been found in animals where adaptation rate is affected by the learned reward value of the odor.

Finally, it should be noted that in some cases, repeated or prolonged exposure to odorants can enhance sensitivity to those odorants rather than induce adaptation. Perhaps the most studied example is sensitivity to the stimulus androstenone, one component of human male sweat and boar saliva. A subset of the human population is anosmic to androstenone, or has very high thresholds for detecting this stimulus. Thus, human sweat smells different to different people. Repeated exposure to androstenone over the course of several weeks, however, can induce or enhance responsiveness. The precise mechanism of the experience-dependent sensitization is unknown, but as with olfactory adaptation may involve both changes in the nose and the brain.

Neurobiology

Olfactory adaptation may involve changes at both the olfactory receptor neurons in the nose and in the olfactory pathway within the central nervous system. Early evidence of central contributions to odor adaptation came from observations that exposure to an odor unilaterally (through one nostril) induces bilateral adaptation. That is, intensity estimates are reduced when that odor is subsequently tested in either the same nostril or the other nostril.

Olfactory receptor neurons can adapt to odor stimulation. How much this receptor adaptation contributes to perceptual adaptation is unclear, however, particularly given the cognitive modulation previously described, which presumably involves central processing. Nonetheless, when odors enter the nose, they bind to olfactory receptor neurons much like a key enters a lock. When the odor key enters the receptor lock, channels in the receptor neuron open and allow charged atoms to enter the neuron. These charged atoms allow the neuron to signal to the brain that a particular odor or piece of an odor has been inhaled. Interestingly, however, these same charged atoms change the proteins within the receptor neuron to reduce how strongly it will respond to the same odor on the next inhalation. This is called *receptor adaptation,* and can last for seconds to several minutes.

Another mechanism that may contribute to odor adaptation relates to odor sampling. Odor molecules are drawn across the olfactory receptor neurons through respiration in humans and terrestrial mammals. Respiration/sampling rate can be switched between a normal respiratory rate of about 2 to 3 sniffs per second to an active sniffing rate of about 5 to 10 per second. Recent evidence suggests that during active sniffing, the olfactory receptor neurons may become less responsive to stable odor inputs and more responsive to dynamic inputs, perhaps via the short-term adaptation of receptors previously mentioned. Thus, as a mouse wanders through a bakery filled with the odor of baked bread, it may come across a piece of cheese. On any single inhalation, both bread odor molecules and cheese odor molecules will be inhaled and activate olfactory receptor neurons. What does the mouse smell? By switching respiration to fast sniffing, the mouse may be able to cause its receptor neurons to adapt to the stable bread background odor more than the just-arrived cheese odor, allowing it to smell the cheese.

There are at least two mechanisms within the central olfactory pathway of the brain that may also contribute to olfactory adaptation. First, olfactory receptor neurons connect to neurons in the olfactory bulb of the brain. Neurons within the olfactory bulb then connect to the olfactory cortex. One population of neurons within the olfactory bulb releases the neurochemical dopamine that acts as a volume control on the input from the olfactory receptors. When very little information is coming into the brain from the receptors, such as when your nose is stuffed up, these neurons reduce the release of dopamine and the volume is increased. This allows the olfactory bulb to become more sensitive to any odors that may make their way up the nose during your cold. If on the other hand, when odor concentrations are high, more of this dopamine could be released, and sensitivity to olfactory nerve input to the brain is decreased. These changes require extended periods of stimulation or deprivation, with total dopamine levels fluctuating over the course of days, rather than minutes, and thus may be more important for long-term adaptation. Interestingly, a very similar process occurs within the eye when we become adapted to the dark or strong light.

As previously mentioned, the olfactory bulb neurons receive direct input from the olfactory receptor neurons and in turn project to the olfactory cortex. In both rodents and humans, it has been found that the olfactory cortex adapts to odors much faster than the olfactory bulb. This adaptation in the olfactory cortex is due to a change in another neurochemical called glutamate. When odors are present for more than a minute or so, glutamate is reduced in the olfactory cortex and the response of the cortex to that odor is reduced. Interestingly, several behavioral measures of odor adaptation correlate strongly with how much the olfactory cortex is adapted. Thus, if the decrease in glutamate can be prevented within the olfactory cortex of a rat, the rat fails to adapt to odors.

The olfactory cortex receives many kinds of information other than what is entering the nose. For example, the olfactory cortex also receives information about whether an organism is awake or asleep, afraid, or hungry. The change in glutamate that occurs in adaptation can be influenced through a biochemical process by these other kinds of information. Therefore, this could provide a mechanism for the attention and/or cognitive control of adaptation previously described. Being told an odor may be dangerous could affect these other inputs to the olfactory cortex and prevent the normal decrease in glutamate from occurring, allowing adaptation to be prevented.

Together, these findings suggest that a number of neural mechanisms may contribute to olfactory perceptual adaptation. The more central processes could help account for variation in adaptation, depending on behavioral state and cognitive processing (e.g., being told an odor may be dangerous). Short-term adaptation may have different underlying mechanisms than long-term adaptation.

Consequences

Olfactory adaptation allows filtering of background or stable odors. This process not only allows us to ignore odors that convey no immediate importance, but may also contribute to background segmentation and analysis of simple odor mixtures into their components. In vision and to some extent hearing, perceptual objects in the foreground can be separated from objects in the background through spatial analysis of the scene. In olfaction, however, spatial analysis is impaired because odors from the background and foreground may be inhaled simultaneously on a given sniff. Despite this merged sampling, it is still possible, for example, to clearly identify the wine at a picnic, despite the various odors of grass and flowers inhaled at the same time. The combination of rapid and highly odor-specific cortical adaptation allows separation of foreground from stable background odors, given at least a brief temporal separation in onset for the background and foreground odors (e.g., your olfactory cortex adapts to the grass and flower odors). Because cortical adaptation is odor specific, however, once the cortex has detected a new odor against the stable background (e.g., the wine), active sniffing may also then contribute to background separation by allowing a suppression of stable olfactory receptor input to the olfactory bulb while maintaining response to new odors. Thus again, adaptation, through a distributed process, contributes to the richness of odor perception.

Donald A. Wilson

See also Context Effects in Perception; Experience-Dependent Plasticity; Olfaction; Perceptual Development: Taste and Olfaction; Perceptual Segregation

Further Readings

Best, A. R., & Wilson, D. A. (2004). Coordinate synaptic mechanisms contributing to olfactory cortical adaptation. *Journal of Neuroscience, 24,* 652–660.

Cain, W. S. (1970). Odor intensity after self-adaptation and cross-adaptation. *Perception and Psychophysics, 7,* 271–275.

Dalton, P. (2000). Psychophysical and behavioral characteristics of olfactory adaptation. *Chemical Senses, 25,* 487–492.

Dalton, P., Wysocki, C. J., Brody, M. J., & Lawley, H. J. (1997). The influence of cognitive bias on the perceived odor, irritation and health symptoms from chemical exposure. *International Archives of Occupational and Environmental Health, 69,* 407–417.

Kadohisa, M., & Wilson, D. A. (2006). Olfactory cortical adaptation facilitates detection of odors against background. *Journal of Neurophysiology, 95,* 1888–1896.

Kurahashi, T., & Menini, A. (1997). Mechanism of odorant adaptation in the olfactory receptor cell. *Nature, 385,* 725–729.

Verhagen, J. V., Wesson, D. W., Netoff, T. I., White, J. A., & Wachowiak, M. (2007). Sniffing controls an adaptive filter of sensory input to the olfactory bulb. *Nature Neuroscience, 10,* 631–639.

Wilson, D. A. (1998). Habituation of odor responses in the rat anterior piriform cortex. *Journal of Neurophysiology, 79,* 1425–1440.

Wilson, D. A., & Stevenson, R. J. (2006). *Learning to smell: Olfactory perception from neurobiology to behavior.* Baltimore: Johns Hopkins University Press.

Zufall, F., & Leinders-Zufall, T. (2000). The cellular and molecular basis of odor adaptation. *Chemical Senses, 25,* 473–481.

OLFACTORY BULB: FUNCTIONAL ARCHITECTURE

The neural architecture of a sensory system reflects two general principles: the physics of stimuli in the relevant sensory modality and the mechanisms and compromises implemented by the system to extract useful information for the organism at minimal metabolic cost. At a high level of abstraction, many of these problems are common to all sensory systems. For example, all

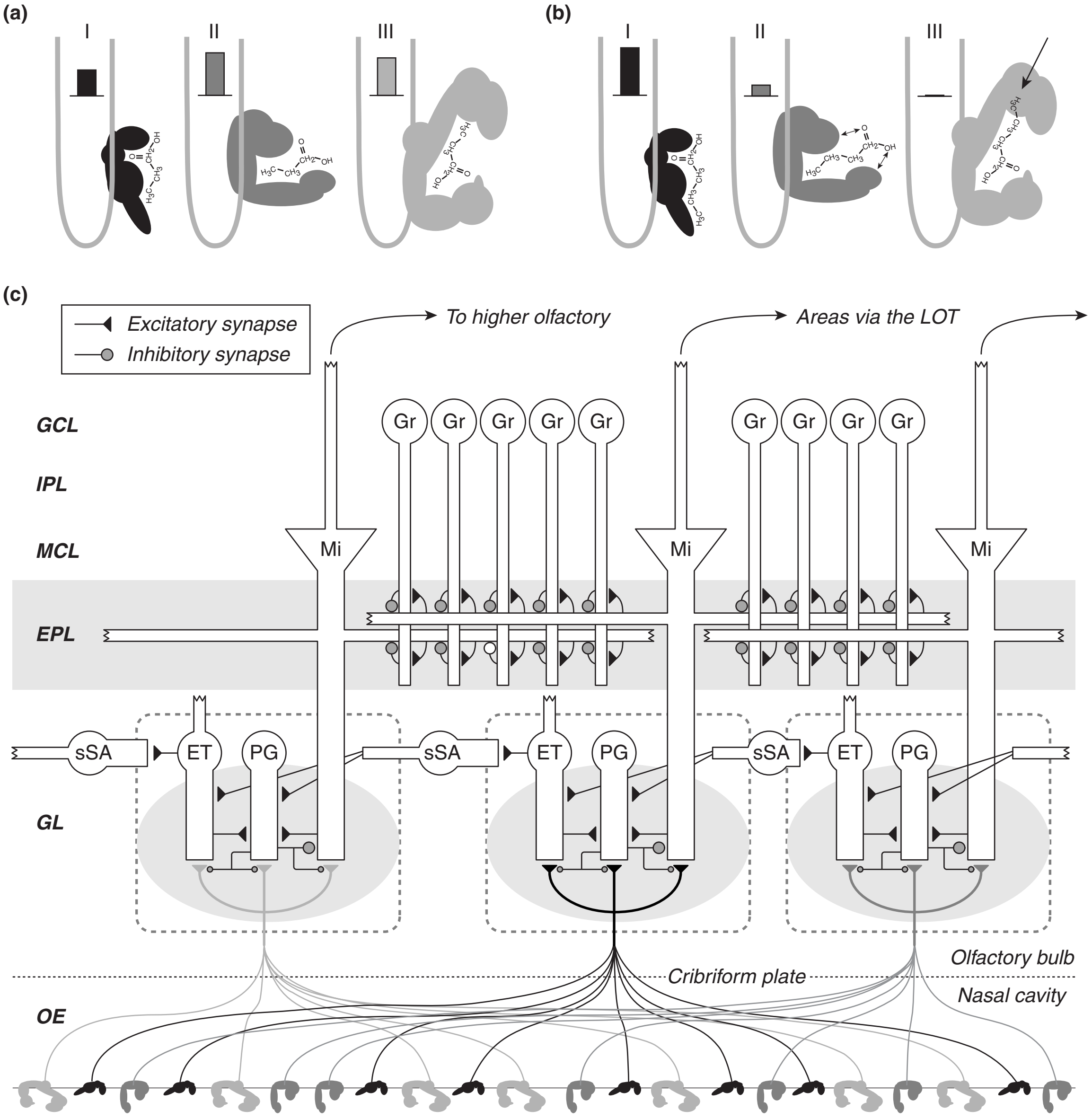


Figure 1 Coding and Anatomy of the Olfactory System

Notes: (a) Illustration of distributed coding in olfactory sensory neurons. Three cilia arising from different OSNs each express a different odorant receptor protein. The three different receptors depicted are each activated to different degrees (vertical bars) by the three-carbon aliphatic odorant propanoic acid. (b) Responses of the same three ORs to the structurally similar four-carbon aliphatic odorant butanoic acid. Receptor I interacts more favorably with this ligand than with propanoic acid, so that the corresponding OSN responds more strongly to butanoic acid at the same concentration. Receptor II, in contrast, does not bind butanoic acid as well as it did propanoic acid, as the polar end of the larger molecule is too far away from its cognate receptor moieties to interact optimally (*arrows*). Nevertheless, butanoic acid weakly activates the receptor. Finally, butanoic acid cannot fit in the binding pocket of Receptor III (*arrow* illustrates impossible fit), and hence does not activate it at all. Reading the relational pattern of activation levels across all three receptors (a three-dimensional vector) enables identification of the odor ligand. (c) Illustration of major olfactory bulb circuit elements. OSNs (represented by their respective OR shapes) are distributed across the nasal olfactory epithelium (OE). The axons of OSNs that express the same OR cross the cribriform plate into the cranial cavity and converge together to form glomeruli (*shaded ovals*) on the surface of the OB. The OSN axon terminals form excitatory synapses (*filled triangles*) onto mitral (Mi), periglomerular (PG), and external tufted (ET) cells. External tufted cells in turn excite PG cells, superficial short axon (sSA) cells, and each other. Periglomerular cells inhibit mitral cell apical dendrites and OSN axon terminals (*open circles*). Mitral cell secondary dendrites extend laterally within the external plexiform layer (EPL, *shaded box*) and form reciprocal synapses with the dendritic spines of inhibitory granule cells (Gr), hence delivering recurrent inhibition onto themselves and lateral inhibition onto other mitral cells. The sSA cells are not affiliated with any given glomerulus, but extend between them, forming a lateral excitatory network consisting of themselves and the ET cells that proliferates in the deep glomerular layer. Shaded ovals connote the approximate physical boundaries of the glomerulus proper, whereas dotted boxes illustrate the group of neurons associated with a particular glomerulus. For visual clarity, only one cell of each type per glomerulus is depicted, and middle/deep (projecting) tufted cells, deep short axon (dSA) cells, and axons arriving from the rest of the brain have been omitted. Other areas shown: LOT = lateral olfactory tract. Olfactory bulb layers, surface to deep: GL = glomerular layer; EPL = external plexiform layer; MCL = mitral cell layer; IPL = internal plexiform layer; GCL = granule cell layer. See also color insert Figure 10.

sensory systems must be able to respond to wide ranges of stimulus qualities and intensities, separate out meaningful stimuli from a background of relatively unimportant stimuli, identify similarities among stimuli so as to be able to classify them accordingly, and learn from experience. The structure of the olfactory system reflects unique functional solutions to each of these common problems. This entry will cover the organizational principles of odor representation and the neural circuits of the olfactory bulb.

Organizational Principles of Odor Representation

Distributed Coding

Primary olfactory sensory neurons (OSNs) are ciliated cells located in an epithelial cell layer lining the nasal cavity (see color insert, Figure 8). Inhaled odorant molecules diffuse through a mucus layer and associate with olfactory receptor proteins (ORs) expressed on these OSN cilia. The OR type that an OSN expresses largely determines its olfactory *receptive field* (also known as its molecular receptive range)—that is, the range of different odorant molecules that will bind to that OR and activate the OSN. Although the OR gene family is extensive—approximately 350 functional genes in humans, over 1,000 in mice and rats—each OSN expresses a very restricted subset of these OR genes; indeed, it is believed that each OSN usually expresses only one of these OR genes. Hence, each OSN is activated by only a subset of all possible inhaled odorants, although any single odorant will generally activate multiple different OSN types to different degrees. This is the fundamental mechanism for *odor representation* in the olfactory system: Any odorant molecule, or combination of molecules, will evoke a unique and characteristic pattern of neuronal activation—a *distributed code*—across the population of OSNs (Figures 1a and b; see also color insert, Figures 10a and b).

This coding scheme has several advantages over a hypothetical alternative in which one OR is specific for each odorant. Whereas the latter scheme could represent 1,000 different odorants in a species expressing 1,000 different OR types, the distributed coding exhibited by the olfactory system can represent far greater numbers of possible odors. Even if OSNs were only counted as "on" or "off," ignoring their capacity to represent intermediate levels of activation, a field of 1,000 ORs could uniquely represent about 10^{300} different odors (compared to an estimated 10^{22} stars in the known universe). Perhaps more importantly, however, distributed coding enables the representation of *similarities* among odors: The more similar two odors are, the more highly overlapping are their neural representations in the nose, as they activate a greater number of ORs in common. This feature enables an animal to either categorize similar odors together (e.g., to identify the smell of "apples" by ignoring the differences in the odors of different apple cultivars or degrees of ripeness) or to discriminate among them depending on motivation and context.

Olfactory Sensory Neurons Convergence

Perhaps the single most spectacular architectural feature of the olfactory system is the OR-selective convergence of OSN axons. The thousands of OSNs that express any one particular OR type are scattered across an animal's nasal epithelium, yet their axons converge together as they project from the nose into the brain—specifically, to the surface layer of the olfactory bulb (OB; Figure 1c; see also color insert, Figures 8 and 10c). In the OB, these converging OSN axons intertwine to form *glomeruli*—clusters of tangled neurites segregated from other glomeruli by a sheath of glial cells. Because of this strict segregation, each glomerulus is generally associated with exactly one OR. This enables odor-specific OR activation patterns to be directly measured via the optical imaging of glomeruli—that is, the activation of a given glomerulus means that the corresponding population of OSNs, expressing a single type of OR, has been stimulated by an odor.

The number of OSNs converging upon a single glomerulus is known as the *convergence ratio,* and it is not uniform among glomeruli. Rather, the ratios reflect functional trade-offs: higher convergence ratios improve the signal-to-noise ratio in OSN populations, increasing sensitivity to low-intensity odors; however, the additional OSNs required for higher convergence ratios consume metabolic energy and space within the olfactory epithelium that otherwise could be utilized by an additional OR type, which would extend the range of detectable odors and/or improve the system's capacity to discriminate among odorants. Indeed, the OSN types that are expressed in a part of the rodent nose that tends to adsorb the most efficiently detectable classes of odor molecules exhibit lower convergence ratios than do OSN types expressed in other parts of the nose, suggesting that the olfactory system is tuned to compensate for the limitations imposed on it by the physics of odor adsorption and the anatomy of the nasal pathways. This also means that behavioral control over sniffing can alter odor representations to a limited extent.

Glomeruli: A Consequence of OR Diversity

Retinotopic mapping is common in several visual areas of the brain, in which adjacent points in space are represented by adjacent visual neurons; similarly, many auditory brain regions exhibit tonotopic mapping, in which neighboring neurons in the brain are responsive to correspondingly similar sound frequencies. Chemotopic mapping within the OB is somewhat more complex. Because of distributed coding, uniquely identifying an odor requires information about its relative activation of hundreds of different OR/OSN types. That is, instead of identifying a pure tone with a single frequency value or a point of light with its two-dimensional X-Y coordinates in space, quantifying a simple odorant stimulus requires a vector with as many dimensions as there are different ORs (Figures 1a and b; see also color insert, Figures 10a and b).

Humans express about 350 different ORs, whereas mice and rats express over 1,000; therefore, quantifying odor quality in these coordinate systems requires a high-dimensional map. Interestingly, it has been shown mathematically that when high-dimensional maps (such as these chemotopic maps of olfactory receptive fields) are projected onto a two-dimensional surface (such as the glomerular layer of the OB), the maps must become fragmented into a clustered, nontopographical organization in which identically tuned neurons can segregate together, but the physical proximity of differently tuned neurons cannot systematically reflect the similarity or overlap of their receptive fields. This precisely describes the organization of the OB glomerular layer.

The sheer diversity of ORs thereby can explain both the existence of discrete glomeruli and why the location of glomeruli within the olfactory bulb does not predict their receptive fields. Furthermore, this nontopographical chemotopic organization fundamentally changes the underlying mechanisms needed to perform certain neural computations in the OB when compared with other sensory systems. Specifically, decorrelation and other computations that rely on neural representations of stimulus similarity cannot utilize spatially localized or center-surround neural mechanisms, such as nearest-neighbor lateral inhibition, instead requiring nontopographical and learning-dependent mechanisms.

Neural Circuits of the Olfactory Bulb

Within OB glomeruli, OSN axons contact the dendrites of two classes of OB principal neurons, mitral and tufted cells. These principal neurons in turn project axons out of the OB to multiple targets in the brain (Figure 1c; see also color insert, Figure 10c). Despite this straightforward synaptic connection, the sensory information carried by OSNs is substantially modified within the OB by descending cortical and neuromodulatory inputs and several classes of intrinsic interneurons that shape the activation of mitral and tufted cells. The resulting transformations of odor representations can be grouped into two stages: (1) glomerular-layer processing that regulates the synaptic transfer of information from OSNs to mitral and tufted

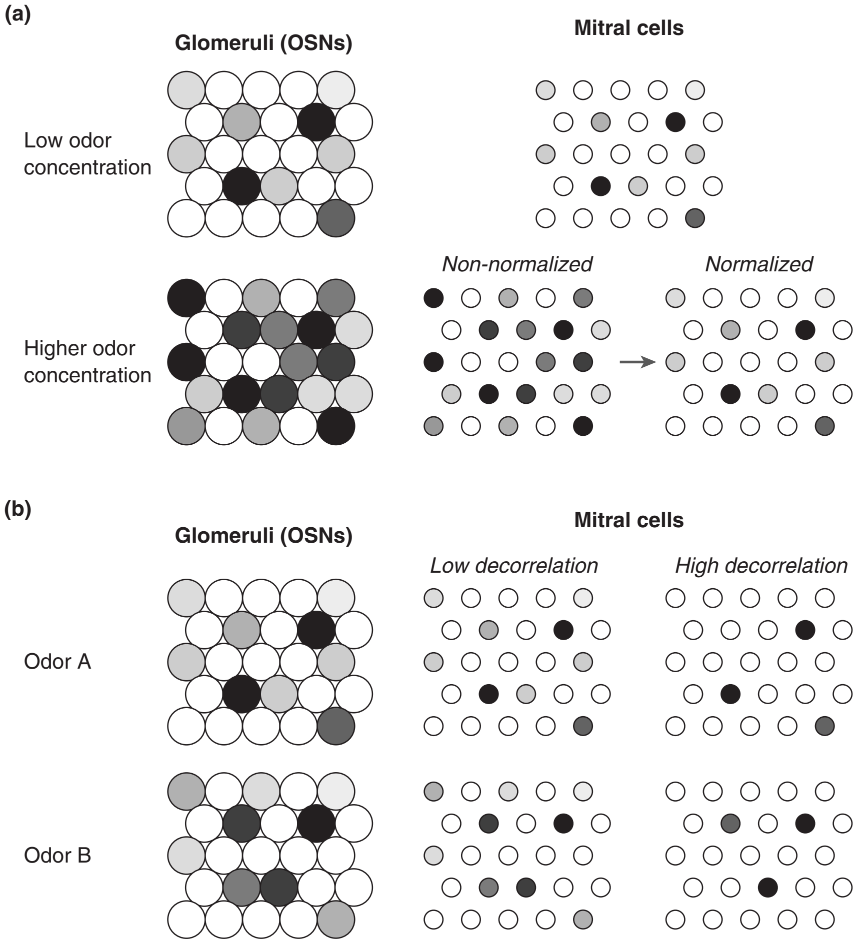


Figure 2 Illustrations of Normalization and Decorrelation in the Olfactory Bulb

Notes: Large circles depict glomerular activation levels that reflect OSN activity patterns; small circles depict the activation levels of the corresponding mitral cells after the effects of OB neural computations have been applied. Darker shades indicate greater activation. (a) Normalization. Higher odor concentrations increase the activity levels of sensitive OSNs and also recruit new, lower-affinity OSNs into the active ensemble. Normalization globally inhibits mitral cells so as to preserve a relational pattern of activation that is less strongly influenced by absolute odor concentration, which may help one recognize the same odor at different intensities. (b) Decorrelation. Odor A and Odor B are relatively similar odors, as evidenced by the substantial overlap in the patterns of glomeruli that they activate. With decorrelation at a minimum, the mitral cell patterns of activation directly reflect these glomerular patterns. At higher decorrelation levels, however, the more weakly activated glomeruli are prevented from activating their corresponding mitral cells, such that the two odor representations become more dissimilar (activating only one mitral cell unit in common).

cells, and (2) subsequent computations in the external plexiform layer (EPL) that further modify the firing patterns of these cells.

Glomerular Layer Circuitry

Besides the axonal arbors of OSNs and the dendritic arbors of mitral (Mi) cells, glomeruli contain dendrites from inhibitory periglomerular (PG) cells and excitatory external tufted cells (ET cells, not to be confused with the projecting type of tufted cell previously discussed). PG neurons inhibit mitral cells; in particular, a subclass of PG cells receives monosynaptic input from OSNs and directly inhibits mitral cell dendrites and OSN presynaptic terminals in a mixed feedforward/feedback loop. ET cells are interconnected by a network of superficial short axon (sSA) cells in a lateral excitatory network that also activates PG cells, thereby inhibiting mitral cells (Figure 1c; see also color insert,

Figure 11c). Interestingly, aside from this broad sSA cell network, and a relatively small number of PG cell axons that project to other glomeruli, the different glomeruli within the same OB are not substantially interconnected.

This glomerular circuitry contributes to *normalization* of the intensity of sensory input (Figure 2a). Whereas collective OSN activity levels can vary over orders of magnitude, average spike frequencies in mitral cells are considerably more constrained, generally being only modestly inhibited or excited by increasing odor concentrations. This probably improves the olfactory system's capacity to recognize the same odor at different intensities, and avoiding high rates of spiking also conserves metabolic energy. Normalization is sometimes confused with *adaptation;* the two processes do interact but are essentially distinct. Glomerular and EPL circuitry are both also credited with mediating *decorrelation* (contrast enhancement) among odors, which regulates the degree to which odors are represented as similar or dissimilar (Figure 2b).

External Plexiform Layer Circuitry

In contrast to the glomerular layer, there is massive interconnection across the OB within the external plexiform layer (EPL). In addition to their dendrites that innervate glomeruli, mitral cells have a second type of dendrite that projects laterally within the EPL and excites the dendrites of granule cell interneurons (Gr; Figure 1c; see also color insert, Figure 10c). The granule cell dendrites in turn reciprocally inhibit the lateral dendrites of the same mitral cell (*recurrent inhibition*) as well as those of other mitral cells (*lateral inhibition*). Lateral inhibitory interactions among mitral cells via intermediating granule cells are thought to mediate the EPL contribution to olfactory decorrelation (Figure 2b). Unlike the retina, however, the strength of this lateral inhibition between two mitral cells is probably unrelated to their proximity. The underlying synapses are plastic, being regulated by inputs received from other regions of the brain and probably by olfactory perceptual learning as well, suggesting that this OB layer contributes heavily to the learning of meaningful odors and mediates some of the effects that such odor memories may have on olfactory processing.

Descending Regulation and Neuromodulation of Olfactory Bulb Circuitry

In addition to afferent olfactory input from OSNs, the OB receives incoming projections from higher olfactory brain areas, as well as neuromodulatory projections from the locus coeruleus (norepinephrine), the horizontal limb of the diagonal band of Broca (acetylcholine), and the raphe nucleus (serotonin). Peptide neuromodulators also are released into the OB both via similar projections and the general circulation. These extrinsic inputs regulate OB physiology and stimulus processing via multiple mechanisms, often with functionally related receptors expressed in different locations within the OB. For example, nicotinic acetylcholine receptors in the glomerular layer and muscarinic acetylcholine receptors in the EPL both respond to cholinergic inputs and together affect OB odor processing, in part by regulating the decorrelation of similar odorant representations (Figure 2b). Indeed, there are more descending projections into the OB than there are afferent projections out of the OB, indicating that the OB is considerably more integrated with higher centers than has often been appreciated.

Thomas A. Cleland

See also Context Effects in Perception; Experience-Dependent Plasticity; Neural Representation/Coding; Olfaction; Olfactory Adaptation; Olfactory Quality; Olfactory Receptors and Transduction; Perceptual Development: Taste and Olfaction

Further Readings

Cleland, T. A., & Linster, C. (2005). Computation in the olfactory system. *Chemical Senses, 30*(9), 801–813.

Cleland, T. A., & Sethupathy, P. (2006). Non-topographical contrast enhancement in the olfactory bulb. *BMC Neuroscience, 7,* 7.

Farbman, A. I. (1992). *Cell biology of olfaction.* Cambridge, UK: Cambridge University Press.

Doty, R. L. (Ed.). (2003). *Handbook of olfaction and gustation* (2nd ed.). New York: Marcel Dekker.

Holscher, C., & Munk, M. (Eds.). (2008). *Mechanisms of information processing in the brain: Encoding of information in neural populations.* Cambridge, UK: Cambridge University Press.

Kohonen, T. (1982). Self-organized formation of topology correct feature maps. *Biological Cybernetics, 43,* 59–69.

Schoenfeld, T. A., & Cleland, T. A. (2005). The anatomical logic of smell. *Trends in Neurosciences, 28*(11), 620–627.

Wilson, D. A., & Stevenson, R. J. (2006). *Learning to smell: Olfactory perception from neurobiology to behavior.* Baltimore: Johns Hopkins University Press.

OLFACTORY CENTRAL PROCESSING

The olfactory processing starts in the periphery at the nasal level where stimuli are presented to the olfactory neuroepithilium and the sensory information is signaled toward the central pathways. Signal transduction leads to a variety of simultaneously occurring events that encode complex olfactory outputs, and the final manipulation of sensory information in the higher brain regions results in olfactory-guided behavior. The neural mechanisms in the human olfactory system have evolved to discriminate a diverse array of odorant molecules initially on the sensory (receptor) level in the nose, and the central olfactory pathways and relay centers further manipulate complex sensory inputs related to odorant molecule recognition and discrimination. The neural machinery that brings about such perception is the topic of this entry.

The Olfactory Receptors and the Olfactory Bulb

Humans have 6 to 10 million olfactory receptor cells located in the upper recesses of the nose. Each of these cells extends 20 to 30 threadlike cilia into the mucus. These cilia contain the G protein receptors to which odorants bind. When an adequate number of receptors are stimulated, an ion flux is induced across the surface of the cell. This flux ultimately results in the release of a chemical "neurotransmitter" at junctions, termed *synapses,* between the activated elongated segment of the cell, termed the *axon,* and cells to which is it is functionally connected.

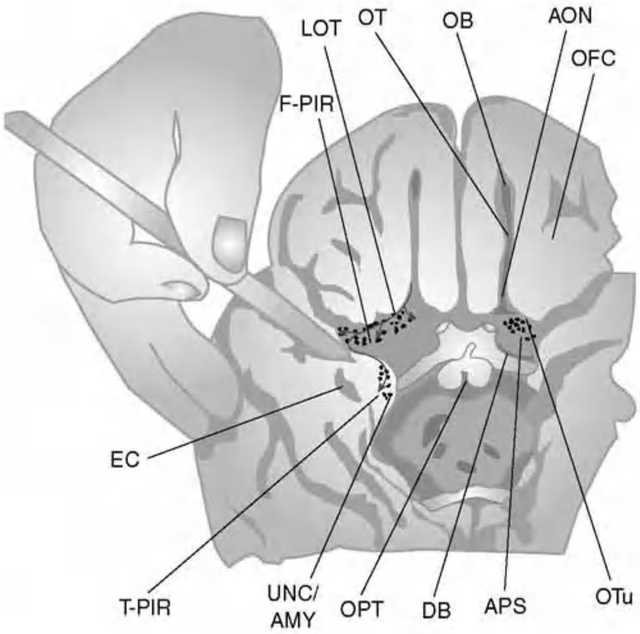


Figure 1 The Anatomy of the Human Basal Forebrain and Medial Temporal Lobes, Including the Olfactory Bulb, Tract, and Surrounding Nonolfactory Structures

Source: Reproduced and modified from Heimer (1983) with permission.

Notes: DB = diagonal band; EC = entorhinal cortex; F-PIR = frontal piriform cortex; OB = olfactory bulb; OPT = optic tract; OT = olfactory tract; OTu = olfactory tubercle; T-PIR = temporal piriform cortex; UNC/AMY = uncus with amygdala situated beneath.

The axons of olfactory receptor cells collect into bundles, termed *fila,* within the deepest levels of the olfactory epithelium and subsequently traverse the bony plate separating the nose from the brain cavity. This "cribriform" plate has many small holes through which these bundles pass. Once inside the brain cavity, the axons disperse from the fila to form the outermost layer of the olfactory bulb, an ovid-shaped structure located at the base of the brain (Figure 1). The recipients of olfactory receptor cell neurotransmitters, most notably the mitral and tufted cells, in turn project their axons to higher brain centers. The mitral cells were named by a famous Spanish anatomist, Ramon y Cajal, because their cell bodies resemble miters, the traditional headdresses of bishops of the Catholic Church.

The initial synaptic interactions between the incoming axons of the olfactory receptor cells and the receiving ends (dendrites) of the mitral and somewhat smaller tufted cells occur within *glomeruli,* globelike structures that make up one of several distinct cellular layers of the olfactory bulb,

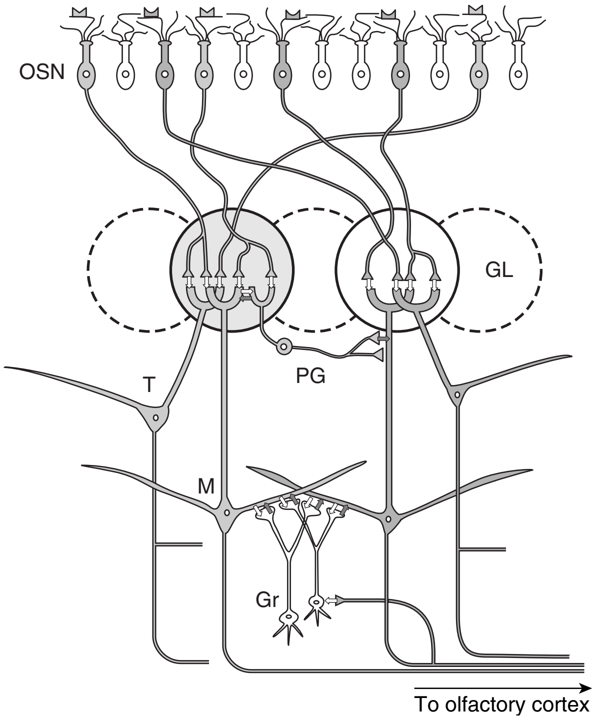


Figure 2 Basic Circuit Diagram Summarizing the Synaptic Organization of the Mammalian Olfactory Bulb

Source: From Mori et al., 1999, with permission.

Notes: Two glomerular modules (gray and white circles) represent two different types of odorant receptors. Mitral cells (M) and tufted cells (T) are output neurons, and granule cells (Gr) and periglomerular cells (PG) are local interneurons. Also shown are the OSN = olfactory sensory neuron, and the GL = glomerulus. Short white arrows denote excitatory synapses, and short black arrows denote inhibitory synapses.

the glomerular layer. The glomeruli are surrounded by cell bodies from some tufted cells and from small cells termed *periglomerular cells*. The latter cells project spiny processes into the glomeruli, where they form two-way (reciprocal) synapses with the dendrites of mitral and tufted cells. These cells also contact dendritic processes of mitral/tufted cells that exit from neighboring glomeruli, affording some degree of control over their signal activity. Cells within the middle-most layer of the olfactory bulb, *the granule cell layer,* also extend processes that interact with mitral/tufted cells, generally inhibiting their activity. Drugs that inhibit granule cell activity produce more robust and generalized activity of the mitral cells, releasing the inhibitory activity. The basic interactions of the receptor, mitral, tufted, and periglomerular cells are shown diagrammatically in Figure 2.

A key element for understanding the first stages of olfactory transduction is that each receptor cell contains only one type of odorant receptor and that each glomerulus receives axons from olfactory receptor cells that express the same type of receptor. Around 400 types of functional receptor proteins are expressed in humans across the 6 to 10 million receptor cells. Because these receptors are tuned to specific molecular features of odorants, the glomeruli are, in effect, functional units, reflecting the precise convergence of common receptor classes. However, the information they receive is complicated by the fact that a given olfactory receptor type can bind a range of molecules that share similar elements, and a given odorant typically activates a number of receptors. Thus, it is the pattern of activation of these receptors from which higher brain regions must extract information related to odorant quality.

Groups of olfactory glomeruli, termed *modules,* are clustered together on the basis of common chemical features that activate the receptors they receive. Chemicals with similar structural features stimulate overlapping with distinct sets of glomeruli within circumscribed regions of the olfactory bulb. Flexible odorants (i.e., ones that can take on multiple configurations and shapes, like hydrocarbons with long straight chains) tend to stimulate more receptors and induce larger regions of glomerular activity than less flexible odorants, such as those made up of short-chained rings. Within a so-called module, the only chemical dimension that has been found to systematically alter general activity is carbon number (or correlated properties such as hydrophobicity or molecular length). Such activity within a module seems relatively insensitive to functional group positions and changes involving the presence, number, position, or stereochemistry of double or triple bonds.

Although glomerular activity increases with odorant concentration, the firing rate of mitral cells is not monotonically associated with such increases, indicating that bulbar output is modulated by inhibitory circuits within the bulb. Such circuits explain, in part, the phenomenon that odorant mixtures are perceived equivalent in intensity to the single odorants they contain.

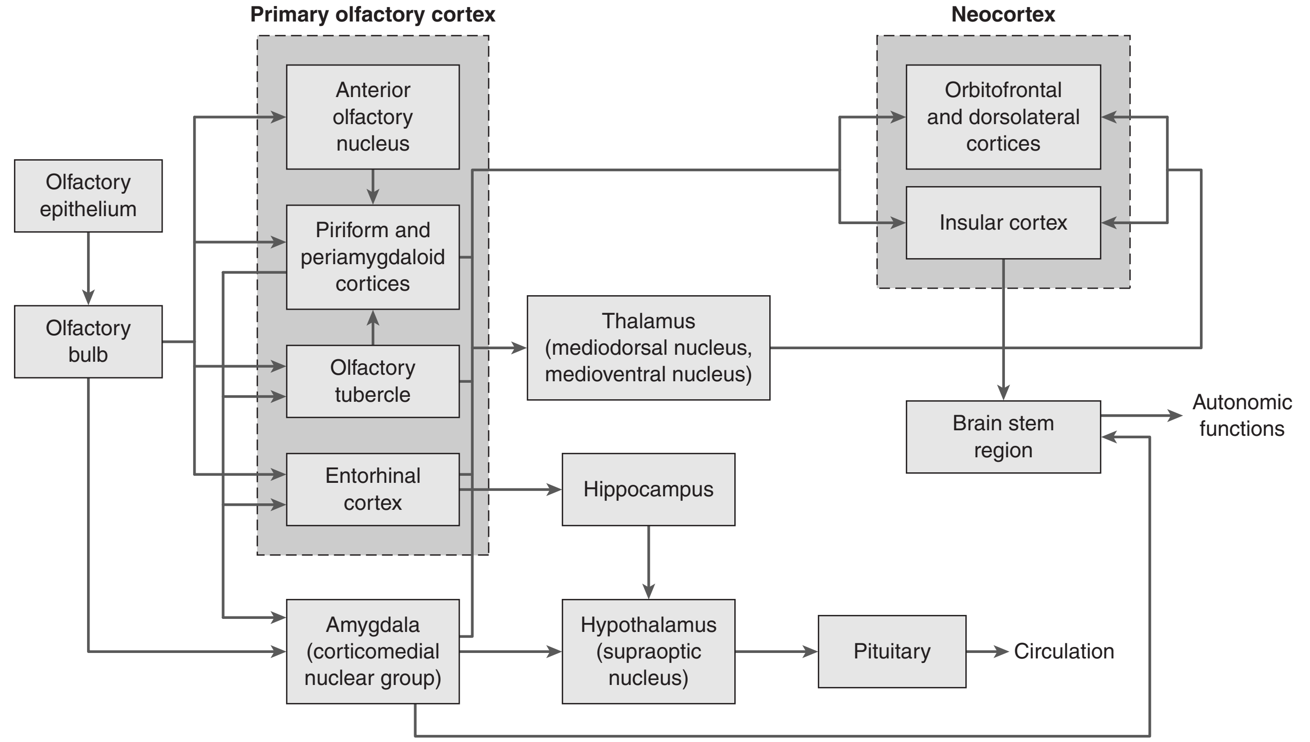


Figure 3 Major Afferent Neural Connections of the Olfactory System in Humans

Source: Copyright © 2008, R. L. Doty.

In general, the size of the olfactory bulbs relative to the size of the brain correlates with the functional importance of olfaction to the organism. For example, the olfactory bulbs of bats that use echolocation to catch flying prey are smaller than the bulbs of fruit-eating bats. In general, nocturnal and ground-dwelling species tend to have relatively larger olfactory bulbs than their diurnal and tree-living counterparts, although bulb size does not always correlate with olfactory sensitivity or function.

The Olfactory Tract and Olfactory Cortex

The primary olfactory cortex is defined as those structures that receive direct projections from the output neurons of the olfactory bulb—the mitral and tufted cells—via the lateral olfactory tract (Figure 3). These structures include the cortical amygdaloid nucleus, the lateral entorhinal cortex, the anterior olfactory nucleus, and the periamygdaloid cortex. The activity of the mitral and tufted cells is modulated by the activity of cells both within and outside of the olfactory bulb. This is in accord with the fact that the olfactory bulbs are not passive relay stations, but modify and filter the incoming signals in accordance with the biological needs of the organism. Rich interactions are known to occur between the elements of the primary olfactory cortex and such important brain regions as the hippocampus, hypothalamus, thalamus, orbitofrontal cortex, agranular insular cortex, and cingulate gyrus.

Anterior Olfactory Nucleus

The olfactory tract makes connections, on its way to higher brain regions, with the anterior olfactory nucleus (AON), a predominantly two-layered cortical-like structure. In humans, the AON has an anterior segment situated in the posterior bulb and a posterior segment located within the anterior temporal lobe. The function of the AON is not well understood, although it aids in the transfer of information passing from bulb to the piriform cortex and vice versa. Importantly, it relays information between the left and right olfactory bulbs via the anterior commissure, one of the nerve fiber bundles that connects the two cerebral hemispheres, and serves a similar role in distributing information to the left and right piriform cortices.

Piriform Cortex

The piriform cortex (PC), named for its pearlike shape, is the largest structure of the primary olfactory cortex and is critical for the processing of basic olfactory information in humans. This structure spans the junction between the frontal and temporal lobes of the brain, having frontal ("prepiriform") and posterior (temporal) components (Figure 3). This three-layered cortex has strong reciprocal connections between a number of brain structures associated with behavior and emotion.

Individual axons from the olfactory bulb mitral cells extend for long distances in the lateral olfactory tract and give off multiple collaterals that can innervate different regions of the PC and the olfactory tubercle. In a manner similar to the mitral/tufted cells in the olfactory bulb, individual pyramidal cells of the PC can be excited and/or inhibited by different odorants. The projection from the bulb to the anterior olfactory nucleus is topographically organized, and the projection to the olfactory tubercle appears to be at least roughly organized along the dorsal-ventral axis. In addition, there is some evidence to suggest that neighboring mitral cells (which often innervate the same glomerulus) project to the same sites in the PC. These findings suggest there may be an underlying organization of inputs in the cortex in which the organization of inputs from the olfactory receptor cells to the olfactory bulb glomeruli is roughly maintained in the projections to the olfactory cortex.

The PC encodes higher-order representations of odor quality, identity, and familiarity, and is intimately associated with the learning and remembering of odors. The frontal component of this cortex preferentially responds to hedonic elements of odors (i.e., pleasantness/unpleasantness), whereas the posterior component does not. The PC plays an important role in the integration of information between olfaction and the other senses. For example, visual stimuli, which under normal circumstances do not elicit piriform activity, come to do so after being paired with a pleasant food odor. The visual response, however, is attenuated when food associated with the odor is eaten to satiety.

Amygdala

The amygdala, which literally means "almond shaped," comprises a group of nuclei within the basal medial temporal lobe that play a key role, along with other elements of the brain's so-called limbic system, in emotion, sexual behavior, and arousal. The corticomedial nuclei of this structure, which receive most of the olfactory bulb input, project to a wide range of brain structures, including the basal ganglia, hypothalamus, orbitofrontal cortex, and thalamus. The amygdala likely plays a significant role in the evaluative (i.e., good/bad) elements of an olfactory perception, as well as in odor memory and its integration with other types of memory. Projections from the olfactory bulb and piriform cortex terminate within sectors of the amygdala, most notably its cortical nucleus. Historically, the amygdala has been viewed as the principal region of the brain involved in odor-related emotional processing with intensity (level of arousal) and valence (degree of pleasantness or unpleasantness) being the fundamental determinants. Recent neuroimaging studies suggest that the amygdala responds only to the intensity of emotionally significant (i.e., pleasant or unpleasant) odors. Its responses to odors of neutral hedonic quality are minimal, even when such odors are intense.

Entorhinal Cortex

The lateral entorhinal cortex is the most caudal of the brain regions that receive axonal projections from the olfactory bulb. The entorhinal cortex has six layers and appears to be the transitional cortex between the so-called three-layered allocortex and the six-layered neocortex. This brain region has a high density of reciprocal projections with the piriform cortex and preprocesses information entering the hippocampus. In conjunction with the hippocampus, it is intimately involved in learning and memory. For example, in nonhumans, it has been implicated in odor matching tasks with delayed intervals, odor discrimination, odor recognition, and trace conditioning. Lesions of the lateral or medial entorhinal cortex disrupt conditioned odor-aversion learning (i.e., avoiding an odor that has been previously associated with sickness or toxicosis).

Orbitofrontal Cortex

The orbitofrontal cortex (OFC) is a five-layered agranular neocortex located in the posterior ventral region of the frontal lobes. This important structure receives input from all primary olfactory regions, including the amygdala, piriform cortex, and entorhinal cortex and, in turn, sends projections back to each of these regions. In addition to rich connections with these structures, the OFC has extensive connections with brain regions subserving taste, vision, and visceral sensations, providing cross-modal integration and associative learning. The OFC presumably facilitates the appreciation of flavor sensations as well as other feeding-related and odor-guided behaviors. Importantly, this olfactory region is intimately associated with judgments of odor familiarity, intensity, hedonicity, and quality.

In aggregate, the OFC is a component of a network of structures involved in decision making, including the short-term maintenance and active manipulation of information during the performance of complex cognitive tasks. The caudal OFC, which is likely the initial neocortical projection from the piriform cortex, has been associated with so-called low-level elements of olfactory processing, such as odor detection. More rostral regions of the OFC are engaged in so-called higher-order olfactory processes, such as working memory, associative learning, and short- and long-term odor recognition memory. The medial OFC and ventromedial prefrontal cortex appear to be more activated by pleasant odors, and the lateral OFC and neighboring inferior prefrontal cortex by unpleasant odors. This medial–lateral responsiveness to pleasant and unpleasant stimuli occurs for other senses as well, implying that these regions are important for general sensory hedonic processing.

Laterality in Olfactory Information Processing

The major olfactory projections to the cortex largely remain confined to the side of the nose from which they arise and do not initially make connections with the thalamus, the great sensory relay station of the brain. This pattern of projections makes it possible to assess, to a large degree, the function of each side of the brain separately by unilateral olfactory stimulation. Thus, odors are more likely to be correctly named when sniffed through the left nostril, tapping the language centers of the left side of the brain. In contrast, odors are found to be more pleasant when sniffed through the right nostril, implying a right-side involvement in odor hedonics.

A number of functional imaging studies suggest that the right OFC is more strongly activated by odors than the left orbitofrontal cortex. In one study employing positron emission tomography (PET), the right amygdala–hippocampal complex was especially activated during recall of emotional autobiographical episodes. In accord with this right-side involvement is evidence that damage to right-sided brain structures, including the amygdala and piriform, entorhinal, and orbitofrontal cortices, has a larger impact on odor discrimination ability than damage to similar left-sided brain structures.

That being said, one cannot escape the fact that some information is transferred to the opposite side of the brain via the anterior commissure. In a classic study, infant rats were conditioned to move toward an odor when one naris was occluded. They were then tested when the previously open nostril was occluded and the occluded nostril was open. When six-day-old rats, who lack a mature anterior commissure, are trained in this way, no transfer of training to the opposite side occurs. However, when 12-day-old rats, who have a mature anterior commissure, are similarly trained and tested, transfer of training does occur to the opposite nostril. Surgical sectioning of the anterior commissure prior to training eliminates such transfer.

Although this entry reflects our growing understanding of the role of each of the central elements of the olfactory system in odor perception, a basic understanding of central odor perception is still lacking. It would seem that the ability to recognize an odor relies upon a number of brain structures that work in concert, including the piriform, entorhinal, and orbitofrontal cortices. Such interconnected structures form nodal points of a “neural network" necessary for the ultimate elicitation of behavioral, cognitive, or emotional responses to odors.

Muhammad Shah, Mudasar Khan, and Richard L. Doty

See also Olfaction; Olfactory Bulb: Functional Architecture

Further Readings

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., et al. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience, 6,* 196–202.

Brewer, W. (Ed.). (2006). *Olfaction and the brain.* Cambridge, UK: Cambridge University Press

Doty, R. L. (Ed.). (2003). *Handbook of olfaction and gustation* (2nd ed.). New York: Marcel Dekker.

Doty, R. L. (2008). The olfactory vector hypothesis of neurodegenerative disease: Is it viable? *Annals of Neurology, 63,* 7–15.

Doty, R. L., Bromley, S. M., Moberg, P. J., & Hummel, T. (1997). Laterality in human nasal chemoreception. In S. Christman (Ed.), *Cerebral asymmetries in sensory and perceptual processing* (pp. 497–542). Amsterdam: Elsevier.

Gottfried, J. A., Deichmann, R., Winston, J. S., & Dolan, R. J. (2002). Functional heterogeneity in human olfactory cortex: An event-related functional magnetic resonance imaging study. *Journal of Neuroscience, 22,* 10819–10828.

Heimer, L. (1983). *The human brain and spinal cord: Functional neuroanatomy and dissection guide.* New York: Springer.

Hudson, R. (1999). From molecule to mind: The role of experience in shaping olfactory function. *Journal of Comparative Physiology A, 185,* 297–304.

Mori, K., Nagao, H., & Yoshihara, Y. (1999). The olfactory bulb: Coding and processing of odor molecule information. *Science, 286,* 711–715.

OLFACTORY IMAGERY

Olfactory imagery is the capacity to experience a smell, intentionally, when the appropriate stimulus is absent (e.g., sniff and imagine the smell of lemon). This is referred to as *phenomenal imagery.* Phenomenal imagery is difficult to study, as it relies upon self-report, which is hard to validate. Consequently, many investigators explore whether imagining an odor has measurable effects on behavior that are akin to those observed when its real equivalent odor is smelled (e.g., does imagining a smell help one to detect its real equivalent?). This is referred to as *performance imagery.* Performance imagery *assumes* the presence of an odor image, but as described in this entry, other causes may produce apparently similar results. Finally, an olfactory image may be evoked automatically and can come to form an indistinguishable part of the percept (e.g., smell vanilla—if you judge this as "sweet" smelling then you have just experienced this form of imagery). In this case, volition has no part to play, there is an odor stimulus, but the resulting percept is a combination of image and stimulus. This is referred to as *cue-driven imagery.*

Phenomenal Imagery

People report that they can imagine odors. However, they also report that their images are not very vivid and are hard to generate, relative to their experience of visual and auditory images. Some more specific findings have also emerged. Olfactory experts report being better able to imagine odors than nonexperts and factor analysis of self-report measures of imagery in all modalities indicates that there is a separate olfactory/gustatory (taste) dimension. That these two modalities should appear together is no surprise, as they routinely combine to form flavor.

The difficulty in generating olfactory images and their lack of vividness may result from a number of causes. It has been suggested that smell perception may be less vivid and distinct than sensory experience in the other modalities, and so this may be reflected in imagined experiences too. However, certain smells can clearly capture attention (e.g., burning or leaking gas) and others may readily elicit strong emotions or distant memories. Consequently, this may not provide a convincing explanation for the reported lack of vividness and difficulty of evocation.

A further reason is that normal participants may simply be unpracticed at forming olfactory images, and the finding that experts are better in this regard suggests this possibility. Similarly, many imagery experiments begin in the following way: “Imagine the smell of *lemon.*” To form the image, some connection must exist between the word *lemon* and a memory of what lemon smells like. Interestingly, normal participants are very poor at naming odors (unlike olfactory experts) that are going from the odor percept to the name. It has been suggested that this is the flipside of going from a name to an odor memory, and that the impoverished nature of the odor-name link—which may result from lack of practice or poor neural connections—may explain why our capacity to imagine odors is not as good as imaging visual or auditory objects. Support for this notion comes from several sources. Hard to name odors are also hard to imagine (unlike their visual equivalents—try to visualize the smell of peanuts versus the sight of peanuts), but learning the odor’s names improves self-report measures of imagery. More interestingly, nearly all of the successful performance imagery studies described in the next section involved participants learning the target odor names prior to imagining them.

Performance Imagery

Three problems make interpretation of the performance imagery literature difficult. The first is that if a perceptual condition (where the odors are really experienced) reveals similar performance to an imagery condition (where the same odors are imagined), the two conditions, real and imagined, will not appear to differ. Unfortunately, this type of outcome, termed a *null result,* is very hard to interpret, as there are almost an infinite number of possible explanations for why something did *not* differ. The second is that many studies do not adequately exclude more compelling alternative explanations for any observed similarity. People may have considerable semantic knowledge about the way odors behave and their properties. Semantic knowledge is usually able to be verbalized and reflects what one knows about a particular thing (e.g., lemons are yellow and taste sour). This semantic knowledge might underpin performance, resulting in equivalent outcomes between real and imaginary experimental conditions. For example, people may know that a stronger odor masks a weaker one, thus, if they are asked to imagine a strong smell while actually sniffing a weak one, they may report that the weak one has been masked. This problem has been effectively overcome in some of the most recent reports considered here. Finally, even if one convincingly demonstrates that performance under imaginary and real conditions is equivalent, and that it does not result from the use of semantic knowledge, how can one be sure that it does result from experiencing an odor image? Attempts to tackle these thorny problems are persistent themes throughout the performance imagery literature.

An impressive variety of procedures have been deployed to examine whether olfactory imagery and perception have similar effects. Two examples illustrate this diversity. First, participants asked to imagine smelling foul odors demonstrate an observably different pattern of sniffing compared to when they imagine pleasant smelling odors. This mirrors the way people behave with real equivalent odors. Second, qualitative similarity (i.e., what something smells like) has been examined for sets of imagined and real odors. There are substantial similarities between real and imagined sets of odors (e.g., lemon and orange odors are judged more alike than boot polish when odors are imagined and when they are smelled). However, as critics such as Hendrick Schifferstein have pointed out, such effects might be obtained by simply asking participants to use their semantic knowledge of odors.

An apparently more definitive finding came from Brian Lyman and Mark McDaniel, who found that the ability to recognize whether one had (or had not) smelled an odor before could be enhanced by imaging its smell. Although others have failed to obtain results akin to this with conceptually similar designs, notably Rachel Herz, the finding that image clarity ratings in Lyman and McDaniel’s study (obtained while participants attempted to imagine the odors) were correlated with recognition memory performance suggests otherwise. It is difficult to see why better self-reported imagery performance should correlate with recognition memory performance if the process was mediated semantically.

The most compelling evidence for performance imagery comes from a recent study conducted by Jelena Djordevic and colleagues. Participants were trained so that they could readily name the two odors used in the study. After establishing participants' threshold for these odors, they were assigned to one of three experimental conditions, olfactory imagery, visual imagery, or a no imagery control. Each group was then given forced-choice detection trials (i.e., is an odor present?). Participants in the imagery conditions were asked to imagine one of the two stimuli just prior to the detection trials. On half of the detection trials, the imagined stimulus corresponded with the target to be detected, and on the other half it did not. They found that detection accuracy was impaired in the olfactory imagery condition, relative to the visual and control conditions, when participants were asked to imagine one odor (e.g., rose) while detecting another (e.g., lemon). These findings are not based upon a null result, as only the olfactory imagery condition affected performance. They also exclude a semantic mediation effect, as participants in the visual imagery condition would have had access to the same semantic knowledge about the odors as participants in the olfactory imagery condition. Finally, they also obtained some evidence for a relationship between participants' self-report of imagery ability and performance on the task—but only in female participants.

A further approach to studying performance imagery has been to establish whether brain areas involved in olfaction are activated when participants imagine odors. While favorable evidence has been obtained by Robert Henkin, the findings are still open to interpretation. In particular, one recent study revealed activation of the primary olfactory cortex simply via reading olfactory-related words (e.g., cinnamon). Participants were never instructed to form images. Thus, the repeated concern about the role of semantic, as opposed to perceptual processes, arises here as well.

Cue-Driven Imagery

Two different forms of cue-driven imagery have been demonstrated. The most compelling example of the first form, where the cue comes from another sensory modality, is from an experiment on wine tasting. George Morrot and colleagues had participants describe the characteristics of white wine and white wine that had been colored red. Participants described the odor of the red-colored white wine in a manner consistent with descriptions of real red wine. A large number of similar findings have also been obtained in which color has striking effects on participants' judgments of odor intensity, naming, quality, and hedonics. One explanation for these effects is that the color cue automatically triggers odor perceptual memories that are consistent with this visual information, and that this top-down perceptual memory overwrites information derived from bottom-up stimulus driven processing. However, evidence favoring this type of explanation has not been forthcoming, and the alternative, in which participants utilize semantic knowledge to generate their evaluations, is better supported. In a recent study, Richard Stevenson had participants engage in an odor discrimination task, where the fluids carrying the odor were incidentally colored. Participants were poorer at discriminating similar odors (e.g., cherry and strawberry) when these were presented in inappropriate colors (e.g., in green fluid). Importantly, this effect was eliminated when participants had to engage in a task to prevent them using verbally based (i.e., semantic) information.

A second form of cue-driven imagery is where the cue comes from the olfactory modality. This can be conceptualized as a form of fragment completion, where the completion appears to draw upon recovering a memory of the whole. The following type of experiment serves as a good example of this. Participants are exposed to an odor-taste mixture such as lychee and sucrose. Later they smell lychee alone and are asked to judge its characteristics. They will typically report lychee as smelling *sweet*—a property of the taste system. This type of finding has been confirmed in several different studies and suggests that what people experience when they smell an odor is often composed of both stimulus-driven and memory-based components. This memory-based component can be conceptualized as an example of cue-driven imagery.

Interpretations

It is undoubtedly harder to imagine the odor of a lemon than it is to imagine the sight of one. Nonetheless, the evidence to date suggests that we do have some capacity to imagine smells. However, a contrarian perspective is also possible, whereby our performance on a real olfactory task can be affected by imagining a smell, but the imaginary smell may not always be consciously experienced. This alternative interpretation would require researchers to ask why certain people can consciously imagine smells and others cannot.

Richard J. Stevenson

See also Auditory Imagery; Cross-Modal Transfer; Hallucinations and Altered Perceptions; Olfaction; Olfactory Imagery; Synesthesia; Taste; Visual Imagery

Further Readings

Djordevic, J., Zatorre, R., Petrides, M., & Jones-Gotman, M. (2004). The mind's nose: Effects of odor and visual imagery on odor detection. *Psychological Science, 15,* 143–148.

Elmes, D. G. (1998). Is there an inner nose? *Chemical Senses, 23,* 443–445.

Herz, R. S. (2000). Verbal coding in olfactory versus nonolfactory cognition. *Memory & Cognition, 28,* 957–964.

Lyman, B., & McDaniel, M. (1990). Memory for odors and odor names: Modalities of elaboration and imagery. *Journal of Experimental Psychology: Learning, Memory & Cognition, 16,* 656–664.

Morrot, G., Brochet, F., & Dubourdieu, D. (2001). The color of odors. *Brain & Language, 79,* 309–320.

Schifferstein, H. (1997). Perceptual and imaginary mixtures in chemosensation. *Journal of Experimental Psychology: Human Perception & Performance, 23,* 278–288.

Stevenson, R., & Case, T. (2005). Olfactory imagery: A review. *Psychonomic Bulletin and Review, 12,* 244–264.

OLFACTORY LOCALIZATION

Animals must find food, find mates, and avoid potential dangers, such as predators, to survive and perpetuate their species. To achieve these goals, the vast majority of animals use their senses to navigate though their environments. Myriad organisms, from bacteria to humans, sense ambient chemicals to help them navigate. Olfaction (smell), the modality that senses volatile chemicals, is especially important for many organisms.

An odorous molecule contains no spatial information. However, as molecules are deposited on surfaces or spread out in air or water, their pattern of deposition over space and time can provide rich information regarding the location of the source that emitted them. The exact information available, and how this information is used, depends on the species and the medium in which the odor is found (for example, a scent trail on the ground versus airborne molecules traveling downwind from a source). In addition, it is becoming increasingly clear that, although animals can locate objects purely through olfactory cues in some situations, animals usually navigate by integrating olfactory information with information from the other senses. Some topics discussed in this entry are: tracking odor trails, the spatial information present in odors dispersed in fluid, tracking using signals from chemicals dispersed in fluid, the integration of information from odor and other senses, and whether humans can track using odors.

Tracking Odor Trails

Chemical stimuli can be deposited on firm surfaces to form trails. As animals pass, they deposit personal odors and can also create disturbance odors, such as volatile compounds released from crushed vegetation or disturbed ground. Various animals use chemical trails to track odor sources. Ants, for example, leave pheromone trails to find their way and to help others from their nest to follow. Snails can follow trails to locate mates, as can snakes. In addition, various predators and parasites use scent trails to track prey and hosts. Dogs provide a familiar example, and some breeds of dogs keep their noses to the ground continuously as they track.

Animals that follow odor trails may track by casting back and forth across a trail. One interesting question is how, upon first encountering an odor trail, an animal determines direction to the source. Odors deposited on the ground dissipate over time, so odors closer to the source will tend to be stronger. Further, because some components of a chemical mixture will dissipate more quickly than others, the balance of mixture components can change over time. How animals make use of these (and possibly other) cues is not fully understood. Some have suggested that ants may determine direction by comparing the information received via separate antennae. Various animals can detect differences in concentration between antennae, including fruit flies. Still, ants would need to be exquisitely sensitive to differences to assess direction in this fashion. Animals might also sample by initially moving up and down a scent trail, as dogs often do. Whatever information dogs use to determine direction, they can do so by sampling as few as five consecutive tracks, or footsteps. From this result, it is estimated that the odor signals in footsteps change enough in about one or two seconds for dogs to determine direction. Interestingly, dogs have more trouble determining direction from a continuous scent trail.

Spatial Information Present in Odors Dispersed in Fluid

The Reynolds number, *Re,* is useful in characterizing general patterns of dispersion. $Re = lU/v$, where v represents the viscosity of the fluid in which the odor is dispersed (water for aquatic animals and air for terrestrial animals), U represents the velocity of the fluid (speed of the water current or wind), and l represents the length scale of interest.

For very low Reynolds values, (i.e., *Re* of less than one), dispersal is smooth and continuous. A good example is a sugar pellet placed in a petri dish of agar. As the sugar molecules diffuse outward from the source, concentration gradually decreases to form a smooth gradient. Under these conditions, which are common for bacteria, concentration decreases monotonically with distance from the source.

For higher Reynolds values, more common for most aquatic and terrestrial animals, dispersal is much more dynamic. Odor molecules spread out downwind (or downstream) in an expanding cloud, called a *plume* in air or a *flume* in water. Within a plume, turbulence creates patches and tendrils of odorized air of varying concentration interspersed with clean air. From the perspective of a sensor placed downwind, a plume would register as variable and intermittent pulses of odor. Time-average concentration still decreases with distance from the odor source. However, the time over which animals would need to integrate to accurately estimate concentration exceeds the time needed for various animals, including some insects and crustaceans, to turn to face an odor source. Because tendrils expand with distance from the source, the frequency and duration of odor pulses over time also contain information.

Tracking Using Signals From Chemicals Dispersed in Fluid

Bacteria and some other simple organisms, which often operate in environments at low Reynolds numbers, display responses to molecules termed *kineses*. In such organisms, the direction of the movement is not specifically directed toward the odor source. However, as the concentration of an attractive odorant increases (as they approach the source), the organism may spend more time in a stationary position or move more slowly. Also, turning may become sharper or more frequent. The net result is that organisms begin to cluster around the attractive odor source over time.

Higher organisms, generally operating in environments at higher Reynolds numbers, tend to display chemically triggered taxes. Unlike a kinesis, a taxis involves movement directed toward an odor source. For example, when many moths and other insects encounter an attractive odor, they move upwind. When many crustaceans and fish encounter an attractive odor, they move upstream. In both cases, the directional information can be obtained from nonolfactory sources, such as somatosensory perception of flow or visual feedback from optic flow.

Male moths, which fly upwind in response to pheromone signals to locate potential mates, are a popular model of tracking in odor plumes (females show similar behavior when tracking food). When a moth contacts a pheromone plume, it will surge upwind in a zigzagging motion known as counterturning. When he loses contact with the plume, he will *cast* back and forth (up and down) perpendicular to the direction of the wind to reacquire the plume. If he fails to reacquire the plume, he may eventually turn and fly downwind. Cockroaches display somewhat similar behavior walking upwind and will begin to circle when odor is removed from the airstream.

Interestingly, most moth species respond to a uniform, unbroken flow of pheromone by initially surging upwind, then casting across the wind as if they have suddenly encountered clean air. To maintain upwind flight, odors must be rapidly pulsed (2–6 pulses per second) to create the kinds of signals moths encounter in natural odor plumes (i.e., pockets and tendrils of odor interspersed with clean air). At relatively low pulse frequencies, moths alternate between upwind surges and crosswind casting. With increasing pulse frequency, as would occur when odor tendrils become more tightly spaced closer to the odor source, crosswind casting decreases and moths fly more directly toward the odor source. These results suggest that moths are adapted to use the fine-grained structure of odor plumes to guide search behavior, a conclusion supported by the fact that receptor neurons in their antennae and cells in their brains can track the rapid changes in odor concentration found in natural plumes.

Other species have adapted to use the fine-grained structure of plumes as well. For example, crayfish can track odor sources more efficiently when the flumes are more complex in structure (this is true even when one controls for the amount of physical turbulence in water). It is thought that many crustaceans, including lobsters, sample fine-grained dynamics of flumes by flicking their antennules to capture filaments and tendrils of odors in the water.

Integration of Information From Odor and Other Senses

As previously indicated, in most cases, animals seem to use information from other sensory systems to guide search behavior (i.e., they will turn upwind or upstream in response to an odor). For example, some sharks will turn toward the strongest current in response to an odor signal, regardless of the location of the odor source. In addition, sharks and other fish can sense vibrations in water, often associated with turbulence caused by moving prey, using a sensory organ called the lateral line (lateral line receptor cells are structurally similar to hair cells in the inner ear). Sharks prefer targets that combine odor and turbulence, and are less efficient at locating odor sources when the lateral line is destroyed. Some species can still find odor sources without a functional lateral line, but cannot do so if they are also deprived of visual cues. Another example of multisensory integration comes from recent work on fruit flies, in which odors seem to modulate the salience of the visual cues by which the animals navigate.

Information from other sensory systems may actually dominate olfactory cues, depending on context and the stage of tracking. Again, sharks provide an interesting example. At close range, sharks can sense electrical signals from prey. When a source of food odor is placed close to electrodes that simulate the electrical field of a fish, sharks hit the electrodes, and will only hit the odor source if the electrodes are turned off. Another example comes from moths that feed on nectar from flowers. Moths track most efficiently when an olfactory stimulus (flower odor) is spatially coincident with a visual stimulus (artificial flower), and less efficiently when the visual cue and olfactory cue are experimentally separated. In addition, moths will prefer the artificial flower to the flower odor when the two are spatially separated. Such behavior makes sense because ambient odors may be present even after a source has been removed. In many cases, chemical cues alert animals to the presence of an odor source and help guide them to an approximate location, whereas input from other sensory modalities help the animal pinpoint the odor source.

Can Humans Track Using Odors?

Even newborns will orient their head toward a preferred odor. For example, infants spend more time facing a pad that carries the sent of their mother as opposed to the scent of a stranger. Further, with some practice, blindfolded adults can follow odor trails in grass by sniffing as they crawl along (though, given the odor concentrations involved, and the speed with which people follow trails, the jobs of tracking hounds remain quite secure).

Odorant concentration should be slightly higher, and the time of arrival slightly earlier, for the nostril closer to the odor source. Could humans use this information to help locate a source? Very little empirical evidence is available. However, despite a few hints to the contrary, the vast majority of work suggests that humans cannot lateralize odors to a particular nostril, even in the extreme case in which one nostril receives a smell and the other nostril receives clean air. Yet, at sufficient concentration, most volatile chemicals stimulate somatosensory nerves in the nose, causing sensations such as the cooling of menthol and the burn of distilled spirits. At concentrations we can feel, we lateralize quite well. Thus, if we are able to locate odor sources based on differences between nostrils, the ability might depend more on somatosensation than olfaction per se.

Other work, using odorants that cause little or no activation of somatosensory nerves, suggests that humans are only able to locate odor sources when they actively turn their heads. This result suggests that humans may locate odor sources by sampling at different locations to determine where a smell is strongest. Thus humans, like other animals, seem to integrate olfactory information with information from other senses to locate odor sources. There has been relatively little systematic research, but odor guided search in humans is likely to be heavily influenced by higher-level cognition. For example, upon catching a whiff of body odor, a person may inspect likely sources, including their armpits and feet, to determine whether or not they are the offender. What one might call hypothesis-driven odor search may not be unique to humans, but probably plays a relatively important role for us.

Paul Wise

See also Action and Vision; Animal Chemical Sensitivity; Auditory Localization: Physiology; Auditory Localization: Psychophysics; Navigation Through Spatial Layout; Olfaction

Further Readings

Chow, D. M., & Frye, M. A. (2008). Context-dependent olfactory enhancement of optomotor flight control in *Drosophila. Current Biology, 18,* 270–275.

Gardiner, J. M., & Atema, J. (2007). Sharks need the lateral line to locate odor sources: Rheotaxis and eddy chemotaxis. *The Journal of Experimental Biology, 210,* 1925–1934.

Goyret, J., Markwell, P. M., & Raguso, R. A. (2008). The effect of decoupling olfactory and visual stimuli on the foraging behavior of *Manduca sexta. The Journal of Experimental Biology, 210,* 1398–1405.

Grasso, F. W. (2001). Invertebrate-inspired sensory-motor systems and autonomous, olfactory-guided exploration. *The Biological Bulletin, 200,* 160–168.

Hepper, P. G., & Wells, D. L. (2005). How many footsteps do dogs need to determine the direction of an odour trail? *Chemical Senses, 30,* 291–298.

Porter, J., Craven, B., Khan, R. M., Chang, S. J., Kang, I., Judkewitz, B., et al. (2007). Mechanisms of scent-tracking in humans. *Nature Neuroscience, 10,* 27–29.

Vickers, N. J. (2006). Winging it: Moth flight behavior and responses of olfactory neurons are shaped by pheromone plume dynamics. *Chemical Senses, 31,* 155–166.

OLFACTORY QUALITY

The difference between two odors is called *olfactory quality,* and it is distinct from other aspects of an odor, such as strength. Imagine that two identical cups sit on a kitchen counter, one containing coffee and one containing tea. Though they may look quite similar, a simple sniff of the substance in each cup would enable an easy distinction between the two warm beverages. There is little chance that the rich aroma of coffee would be mistaken for the fresh fragrance of tea. One might be described as slightly "nutty," or the other as "floral," but whatever the description, determining that they are different is straightforward.

Historically, attempts have been made to divide odorants according to quality into a few categories, such as fragrant, spicy, or burned, with the idea that these categories might be combined to produce the full range of odorant qualities. The reasoning was that the olfactory system might function similarly to other sensory systems, such as color vision, in which input from three receptors combine to give a full color experience. However, once some of the biological aspects of the olfactory system were uncovered, it became clear that because of the extremely large number of olfactory receptors, a model with a small number of categories would not be able to accurately represent either odorant quality or the qualitative similarities between odorants. Qualitative similarity can be somewhat described using multidimensional scaling (MDS), a mathematical technique that arranges odorants in a geometric space based on qualitative relationships. In an MDS space, more similarly perceived odors (such as sage and thyme) are closer to each other than those that are not (such as coffee and tea). However, this mathematical technique is still heavily based on the particular set of odorants used to create the dimensional space, and thus is restricted to illuminating relationships between odorants rather than describing olfactory quality per se. So, MDS would be able to give an idea of how similar the odor of coffee is to that of tea, but would not address any of the specific qualitative features that make one substance smell uniquely like "coffee."

Unlike the wavelength of light, which is highly correlated with the perception of color, the physical correlates of odor quality are neither easily measurable nor easily described. The chemical composition of an odor can be analyzed with a gas chromatograph, but the human perception of quality is only partially dependent on the chemical structure of an odor. Instead, it is a combination of the chemical structure and the changes that are produced in both the nose and the central nervous system. So, three different factors discussed in this entry, chemical structure, physiology, and cognition, play a role in the perception of odorant quality.

Chemical Factors

The stimuli perceived by the sense of smell are volatile molecules arising from the source of the odor. Members of a large family of odorant receptor proteins recognize features of the molecules, though the specific nature of those features is not as yet clear. Nevertheless, people are able to combine and use those features to distinguish between thousands of different odor qualities. Some evidence suggests that as few as 18 physiochemical features may account for the salient properties that allow odorant discriminations, but there may be many more. One such physiochemical feature may be the length of a molecule's carbon chain, because it is possible for people to discriminate reasonably well between pure odorant molecules in a homologous series (such as aliphatic alcohols, aldehydes, or carboxylic acids) that only differ in terms of carbon chain length. This suggests that carbon chain length may be an important molecular property affecting odor quality, though it is unlikely to be the only aspect of chemical structure that influences quality perception.

Although chemical structure is important, olfactory quality perception is not necessarily related to the chemical composition of an odorant. Structurally related odorants may smell quite differently from each other, and odorants that are unrelated may smell fairly similar. Enantiomers, for example, are molecules that are identical in all ways except their rotation. For example, the chemicals *1*-carvone and *d*-carvone are enantiomers; they are identical, except that they are essentially mirror images of each other. Yet, these two chemicals differ dramatically in terms of their quality: One smells like mint, whereas the other smells like rye. Enantiomers with different qualities illustrate that chemical structure alone cannot account for odorant quality perception. Further evidence that odor quality is not merely due to chemistry comes from the observation that different brain areas are responsible for decoding the chemical structure of an odorant as opposed to its quality. Functional imaging data indicates that posterior regions of the piriform cortex encode quality (so the experience of mint and rye, but not the chemical structure of their molecules), whereas anterior regions encode chemical structure (but not quality).

When discussing the physiochemical properties of odors, it is worth noting that typical, everyday smells, such as coffee or tea, are not pure chemicals, but are instead mixtures made from many different molecules. These mixtures are generally not perceived as such, because mixtures containing more than three components are usually perceived as a single unit or perceptual whole by the olfactory system. So, the perception of odorants is a synthetic process, in which molecules are often combined for quality perception. However, this perceptual system must also be analytic, as these everyday molecular mixtures of volatile chemicals are generally found in an environment that may contain other odorous substances. So, the air in the kitchen may contain a bacon aroma in addition to the smell of coffee and tea. In order to perceive a particular odor quality, the odor must be discriminated from the background. In effect, the coffee must be perceived against the background of the bacon smell, because a single sniff may take in both types of molecular mixtures to the olfactory receptors. This figure-and-ground discrimination is not a particularly easy task, which is the reason that odors are sometimes called inherently noisy stimuli.

But the amazing thing is that it is accomplished at all, because despite all of these different component molecules being mixed together in the nose, we can still discriminate bacon and coffee.

Physiological Factors

The first parts of the human body to detect an odor molecule are called the *olfactory receptors,* and they are located at the very top of the nose in the nasal epithelium. Odorous molecules bind to receptor sites using a "lock and key" mechanism, and when the two fit together, a neural signal is initiated. Because several different receptors will respond to a single odorous molecule, it is important to note that some receptor types respond to the molecule better than others. Further, the range of odorants that will effectively activate a single receptor is quite large. So, some receptors will be better activated by the coffee smell than others, but those receptors may also be somewhat stimulated by the odor of tea. Consequently, in order to determine odor quality, a combinatorial code that takes into account the pattern of responses from each of the receptors is important. In other words, the overall pattern of receptor activity in response to an odor such as coffee is important to perceiving odor quality. The differences in responsivity of receptors lead to a topographical representation of an odorant across the nasal epithelium that is an initial spatial pattern of neural activity uniquely representing the odor. An odor's unique pattern of activity is maintained as the information travels from the epithelium to the olfactory bulb in the brain, and into the piriform cortex. After this point, the highly ordered topographical representation is either lost or becomes very subtle as the information travels to higher brain regions concerned with olfaction. Many areas of the olfactory cortex also send feedback projections to the olfactory bulb that likely provide the pathways for cognitive factors to influence the perception of odorant quality.

Cognitive Factors

Cognitive factors, such as the context in which an odor is perceived, its name, or a memory, may define an olfactory perception as much as a molecular structure or receptor physiology. These factors are likely effective perceptual cues because prior experiences lead to expectations of odorant quality.

Many of the contexts in which an odor is experienced can influence the perception of its quality, including tastes, colors, and emotional states. Odorants that are presented in the context of a sweet taste (like sucrose) are perceived as sweeter than they were before the pairing. Sour tastes can also alter the reported quality of an odorant such that it is rated as more sour smelling after the association. Color is another important part of context for the perception of odorants. Odor quality is more easily discriminated when the odors are colored "appropriately." In other words, people are better at saying that strawberry and cherry smell differently when they are presented in red water than in green water. The color seems to act as a confirmatory cue to the concept of the odor. In another example of the contextual importance of color, wine experts who are given white wine that appears red because of food coloring tend to use words that are normally associated with red wine to describe the flavor. Emotional state can also form an important part of the context for perceiving odor quality. Beliefs about odorants can influence their perception, such that simply telling someone that an odor is healthy or hazardous can bias quality perception positively or negatively. Further, inducing an emotion, such as the experience of failure, in the presence of an ambient novel odorant can cause the odor to be reported as having negative qualities, whereas eliciting success can cause the odorant to be positively perceived.

Although the quality of an odor is difficult to describe, a number of aspects of the perceptual quality of odorants, such as ratings of odorant pleasantness or familiarity, are enhanced when human subjects either can identify the odor or are provided with its name. If people label an odor incorrectly initially, their memory will be related to the label so much that they will mistakenly identify an odor corresponding to their incorrect label as having been previously presented. Further, people will report qualities consistent with the name that they have been given. So, the same odorant (mixture of isovaleric and butyric acids) will be described negatively if presented to the participant as "vomit," but described positively if it is presented as "parmesan cheese."

It has been suggested that the perception of odorant quality is heavily dependent upon the memory of prior experience with the odorant. Experience with odorants that are quite similar to each other allows them to be more easily discriminated. So, the wine connoisseur has little trouble discriminating between the odors of two bottles of cabernet sauvignons. Further evidence comes from an amnesic patient called H. M., who is severely impaired on tasks that involve formation of new long-term memories but performs normally on tasks that rely mainly upon immediate memory. When given a battery of tests for olfactory function, H. M. demonstrated a normal ability to detect odors and discriminate intensity, but his ability to discriminate odor quality was severely impaired. This is in keeping with findings from other patients with temporal lobe damage who also demonstrate a decreased ability to discriminate odor quality, and suggests that memory is an important part of the process of quality perception.

So understanding the human ability to distinguish olfactory quality requires more than simply knowing the chemical composition of an odorant, and more than a simple interaction between receptors and features of molecules. Effective perception of odorant quality also requires input from higher cognitive brain areas, and an ability to make use of information that comes from a level of previous experience with the odors. In short, in a single sniff, a chemical, the body, and past experiences interact for the experience of odor quality, and allow an easy distinction between the smell of coffee and that of tea.

Theresa L. White

See also Context Effects in Perception; Experience-Dependent Plasticity; Olfaction; Perceptual Development: Taste and Olfaction; Perceptual Segregation

Further Readings

Buck, L. M., & Axel, R. M. (1991). A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell, 65,* 175–187.

Cain, W. S. (1979). To know with the nose: Keys to odor identification. *Science, 203,* 467–469.

Chu, S. (2008). Olfactory conditioning of positive performance in humans. *Chemical Senses, 33,* 65–71.

Dalton, P., Wysocki, C. J., Brody, M. J., & Lawley, H. J. (1997). The influence of cognitive bias on the perceived odor, irritation and health symptoms from chemical exposure. *International Archives of Occupational and Environmental Health, 69,* 407–417.

Distel, H., & Hudson, R. (2001). Judgment of odor intensity is influenced by subjects' knowledge of the odor source. *Chemical Senses, 26*(3), 247–251.

Epple, G., & Herz, R. S. (1999). Ambient odors associated to failure influence cognitive performance in children. *Developmental Psychobiology, 35,* 103–107.

Gottfried, J. A., Winston, J. S., & Dolan, R. J. (2006). Dissociable codes of odor quality and odorant structure in human piriform cortex. *Neuron, 49*(3), 467–479.

Herz, R. S., & von Clef, J. (2001). The influence of verbal labeling on the perception of odors: Evidence for olfactory illusions. *Perception, 30*(3), 381–391.

Jacek, S., Stevenson, R. J., & Miller, L. A. (2007). Olfactory dysfunction in temporal lobe epilepsy: A case of ictus-related parosmia. *Epilepsy & Behavior, 11*(3), 466–470.

Laska, M., & Hübener, F. (2001). Olfactory discrimination ability for homologous series of aliphatic ketones and acetic esters. *Behavioral Brain Research, 119*(2), 193–201.

Pollak, E. H. (1973). Multiple profile-multiple receptor site model for vertebrate olfaction. *Journal of Theoretical Biology, 245,* 175–184.

Stevenson, R. J., Boakes, R. A., & Prescott, J. (1998). Changes in odor sweetness resulting from implicit learning of a simultaneous odor-sweetness association: An example of learned synesthesia. *Learning and Motivation, 29,* 113–132.

Wilson, D. A., & Stevenson, R. J. (2003). The fundamental role of memory in olfactory perception. *TRENDS in Neuroscience, 26,* 243–247.

OLFACTORY RECEPTORS AND TRANSDUCTION

Like all sensory stimuli, odors carry information. Odors can tell you whether your food is fresh or spoiled; can warn of dangers like fire, predators, or infection; and may even provide information about the suitability of a mate. Though some odors may elicit the same stereotyped behaviors or

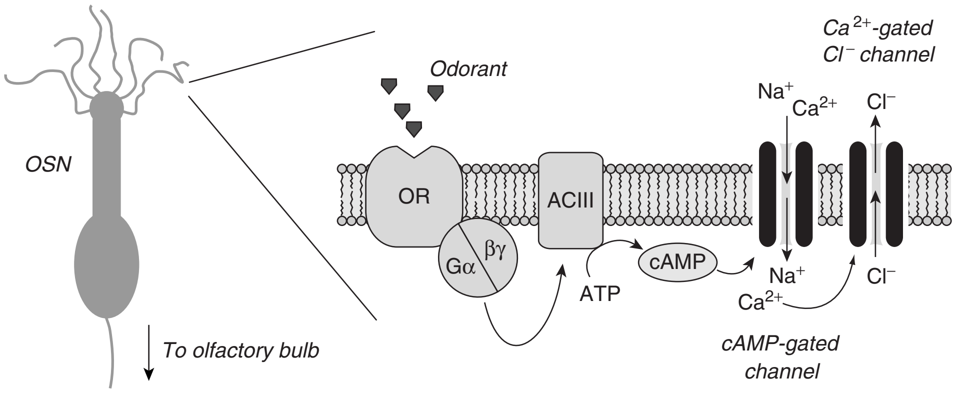


Figure 1 The Olfactory Transduction Cascade in Canonical Olfactory Sensory Neurons (OSNs)

hormonal changes in every individual, the meaning of most odors is heavily dependent on our previous experiences, our motivations and preferences, and the context in which they are encountered. However, no matter what their meaning might be, odors can only be perceived if they are detected. Odors are complex mixtures of chemical stimuli. Each component of a particular odor mixture is termed an odorant. Humans and other mammals use a large repertoire of receptor proteins on the surface of sensory neurons in the nose to recognize distinct odorants and other chemosensory stimuli. Once activated by their interactions with these stimuli, the receptors trigger a biochemical cascade within the cell to transduce, or change, the chemical signal into a neural signal that can be recognized by the brain.

The mammalian nose employs many hundreds of different chemosensory receptors—approximately 350 in humans and close to 1,500 in mice—as well as several distinct biochemical mechanisms to detect and transduce olfactory signals. The large array of receptors increases the number and type of chemicals that can be detected by the olfactory system; the need for several different transduction mechanisms is less clear. Recognition of this receptor diversity has contributed to our realization that the olfactory system is organized into several subsystems, each using a different combination of receptors and transduction mechanisms to recognize and encode olfactory stimuli. Thus, each subsystem is selective for those stimuli that can activate the receptors it expresses.

Why olfactory transduction is so complex remains unclear. However, a complex sensory world may require a complex sensory system. The large variety of chemosensory receptors can expand the universe of potential olfactory stimuli, not only to chemically diverse volatile molecules but to peptides and proteins as well. Distinct, multistep transduction cascades offer numerous opportunities for signal amplification, temporal control, and contextual modulation (e.g., by adjusting olfactory sensitivity in the context of metabolic needs or mating status). Finally, the coexpression of certain receptors, signaling molecules and ion channels, may reflect their own functional limitations such that certain types of receptors can only work efficiently with certain signaling molecules. This entry discusses receptors in the main and accessory olfactory systems.

Main Olfactory System

The main olfactory system is composed of three principal anatomical divisions: (1) the main olfactory epithelium, which contains the olfactory sensory neurons (OSNs); (2) the main olfactory bulb, which is the central nervous system target of the OSNs; and (3) higher olfactory areas such as the olfactory cortex, which receive inputs from the main olfactory bulb (see Olfactory Bulb: Functional Architecture Figure 1; see also color insert, Figure 10). The olfactory regions of the brain are critical for processing odor information received from the periphery, helping to make sense of the signals the nose detects. The main olfactory epithelium, in contrast, is primarily concerned with the detection of olfactory stimuli. When odorants and other chemical stimuli enter the nose, usually with sniffing or during the chewing of food, they diffuse through a thin layer of mucus to reach the olfactory epithelium. These odorants can then interact with odor receptors on the apical tips of OSNs to initiate the process of olfactory transduction.

The number of subsystems within the main olfactory system is unclear. However, those identified so far each use a strikingly similar collection of molecules and mechanisms to recognize and transduce olfactory stimuli: (1) a receptor protein that sits within the plasma membrane of the OSN such that it can interact with both external stimuli (e.g., an odorant) and the transduction cascade within the cell; (2) an enzyme that produces a diffusible molecule (called a second messenger) upon receptor activation; (3) an ion channel that opens in the presence of the second messenger or related intracellular signals (e.g., Figure 1). The culmination of these cascades is the net movement of positive charge into the OSN, depolarizing the cell membrane, and promoting the generation of action potential signals to the brain.

The largest and best understood subsystem of the main olfactory system is known as the *canonical* main olfactory system. Canonical OSNs were the first to be characterized and comprise the largest subpopulation of OSNs in the main olfactory system. They can be most easily distinguished from other OSNs by their expression of a family of odorant receptors known as olfactory receptors (ORs) (Figure 1). These receptors, discovered by Linda Buck and Richard Axel, are members of the G protein coupled receptor (GPCR) superfamily. As such, they share a number of structural and functional attributes with other GPCRs: these include seven-plasma-membrane-spanning domains and the ability to couple to the transducing proteins called G proteins. An individual canonical OSN appears to express only one of the many hundreds of OR genes. Each OR has a distinct selectivity for a small number of odorants, though usually more than one OR can be activated by the same odorant. Therefore, the presence of any one odorant results in the activation of several distinct groups of canonical OSNs. Those OSNs express the same OR project axons to common targets, called *glomeruli,* in the olfactory bulb. The convergence of functionally homogeneous OSNs onto just a few olfactory bulb glomeruli results in the establishment of a distinct pattern of activity in the brain that reflects all of the ORs activated by each odorant.

Although animals such as mice have many more functional OR genes than do humans (more than 1,000 vs. approximately 350), it is believed that most mammals are sensitive to a very similar repertoire of odorants (pheromones and other species-specific olfactory cues are another matter). This similarity may be a consequence of the ability of individual ORs to recognize multiple odorants. The relatively small number of human ORs compared to mouse ORs may reflect a greater degree of olfactory redundancy in rodents and other animals that rely heavily on their olfactory system to sense the environment.

Individual OR genes commonly exhibit variations in their genetic sequence. Variants of the same OR gene are called alleles. Two individuals can have distinct repertoires of OR alleles. Different OR alleles can show functional differences, possibly altering the odorants to which an individual is sensitive. Specific anosmias, where individuals are insensitive to particular odorants but otherwise have a functioning olfactory system, are quite common. Indeed, it is likely that every person has one or more specific anosmias that results from having particular OR alleles. For example, allelic differences in a single OR account for a large percentage of the human variation in olfactory sensitivity to the odorant androstenone.

Olfactory transduction requires more than just receptor activation: Enzymatic production of a diffusible messenger molecule (the second messenger) and the opening of ion channels are required for the generation of action potentials that can carry olfactory information to the brain. In the first step, odorant interaction with an OR results in a change in the shape of the receptor, called a conformational change. This odorant-induced conformational change is transferred through the cell membrane to the cytoplasmic surface, where it stimulates a G protein complex composed of α, β, and γ subunits (Figure 1). The α subunit in canonical OSNs is known as $G\alpha_{olf}$. OR activation promotes the separation of $G\alpha_{olf}$ from the β and γ subunits. The solitary $G\alpha_{olf}$ interacts with another integral membrane protein called Type III adenylyl cyclase (ACIII; Figure 1). One of several isoforms of adenylyl cyclase found in mammals, ACIII acts as an effector of OR activation. Upon interaction with $G\alpha_{olf}$, ACIII synthesizes the soluble molecule cyclic 3'5'-adenosine monophosphate (cAMP; Figure 1), which can freely diffuse through the cell to activate its own targets.

The principal target of cAMP is an ion channel (Figure 1). This channel is similar to those that play an important role in phototransduction in the mammalian retina. The olfactory channel opens when it comes in contact with cAMP, allowing the flux of sodium and calcium ions into the cell. This transfer of positive charge reduces the voltage differential across the plasma membrane. This membrane depolarization promotes the generation of action potentials down the axon to the olfactory bulb, thus signaling the brain that an odorant has been detected. However, the ions entering through this cAMP-gated channel are not wholly responsible for the odor-dependent depolarization of the OSN membrane. Calcium ions entering through the channel also promote the opening of a chloride channel (Figure 1). Outward flow of negatively charged chloride ions increases the net positive charge within the cell, further depolarizing the OSN and increasing the chance that it will reach the threshold for action potential generation.

Calcium ions also play a key role in odor adaptation in canonical OSNs. Odor adaptation is the reduction in odor response upon prolonged or repeated odor stimulation. In other words, with prolonged or repeated odor exposures, a greater amount of the stimulus will be needed to elicit the same amount of response. Calcium ions contribute to odor adaptation by working with the calcium-binding protein calmodulin to desensitize the cyclic nucleotide-gated channel to activation by the second messenger cAMP. Phosphodiesterases, which degrade cyclic nucleotides and thus lower the levels of cAMP in the cell, may also play important roles in odor adaptation and/or the termination of the transduction signal.

Other subsystems of the main olfactory system draw on distinct receptors and transduction mechanisms, and perhaps include both G protein dependent and independent signaling cascades. For example, OSNs in one subsystem utilize a guanylyl cyclase called GC-D to transduce olfactory stimuli, including some peptides, present in urine. The chemosensory receptor in these OSNs has not been identified. However, GC-D itself is a good candidate, as it is a member of a family of peptide receptors. Stimulation of GC-D-expressing OSNs results in an elevation of intracellular cGMP, which in turn opens a cGMP-sensitive channel. As in canonical OSNs, opening of this channel promotes action potential signaling to the brain. Characterization of the various receptors and transduction mechanisms at work in the main olfactory system, as well as the roles of these subsystems in humans, are major areas of ongoing investigation.

Accessory Olfactory System

The accessory olfactory system is composed of three principal anatomical divisions: (1) the vomeronasal organ, which contains the vomeronasal sensory neurons (VSNs) of the accessory system; (2) the accessory olfactory bulb, which is the central nervous system target of the VSNs; and (3) higher olfactory areas, such as the amygdala, which receive inputs from the accessory olfactory bulb. The regions of the brain targeted by the accessory olfactory system impact stereotypical behaviors, such as mating, and can influence hormonal changes. The vomeronasal epithelium is similar to the main olfactory epithelium in that it is the site of chemosensory transduction. In contrast to the main olfactory system, stimuli must be actively pumped into the vomeronasal organ upon direct contact with the stimulus source (often urine or glandular secretions) so that they can gain access to the VSNs. The accessory olfactory system is prominent in most mammals, but appears to have been lost in the primate line contemporaneously with the emergence of trichromatic vision.

Like OSNs in the main olfactory system, VSNs use receptors, second messengers, and ion channels to transduce their signals. However, the precise functions of the receptors, signaling molecules, and ion channels implicated in VSN chemosensory transduction remain poorly defined, at least in comparison with their equivalents in the canonical OSNs. Two subsystems have been identified in the accessory olfactory system. They can be distinguished by the location of their VSNs (apical vs. basal layer of the epithelium), the receptors and G proteins they express, and their targets (anterior vs. posterior) in the accessory olfactory bulb.

The apical division of the vomeronasal epithelium has been implicated in the detection of small, volatile stimuli, many of which are suggested to act as pheromones. Pheromones are olfactory stimuli given off by one member of a species that elicit a behavioral or hormonal change in another member of that same species. The VSNs of the apical division express members of a large family of GPCRs, the type 1 vomeronasal receptors (V1Rs), which likely act as chemosensory receptors. Though somewhat similar in structure to the ORs, the V1Rs are not closely related to ORs. The apical VSNs are also distinguished by their expression of the G protein subunit $G\alpha_{i2}$. Unlike the $G\alpha_{olf}$ protein in OSNs, $G\alpha_{i2}$ has not been definitively linked to chemosensory transduction in VSNs. The basal division of the vomeronasal epithelium may be more specialized for peptide and protein stimuli present in glandular secretions and in urine. The VSNs in this division express another GPCR family, the type 2 vomeronasal receptors (V2Rs), which are unrelated to either ORs or V1Rs but which likely function as chemosensory receptors. Basal VSNs are expressed $G\alpha_o$, though again this G protein has not been definitively linked to chemosensory transduction. The two groups of VSNs appear to use a similar biochemical cascade to transduce chemosensory signals: Activation of the effector enzyme phospholipase C, production of lipid-derived second messengers diacylglycerol and inositol 1,4,5-trisphosphate (IP3), opening of an ion channel that likely includes the diacylglycerol-sensitive channel Trpc2, and the movement of positive charge into the cell. However, the precise roles of all these molecular players, especially their interactions with each other, are yet to be elucidated.

Steven D. Munger

See also Olfaction; Olfaction and Reproductive Behavior; Olfactory Adaptation; Olfactory Stimulus; Taste Receptors and Transduction; Vomeronasal System

Further Readings

Ache, B. W., & Young, J. M. (2005). Olfaction: Diverse species, conserved principles. *Neuron, 48,* 417–430.

Brennan, P. A., & Zufall, F. (2006). Pheromonal communication in vertebrates. *Nature, 444,* 308–315.

Buck, L., & Axel, R. (1991). A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell, 65,* 175–187.

Fain, G. L. (2003). *Sensory transduction.* Sunderland, MA: Sinauer.

Firestein, S. (2001). How the olfactory system makes sense of scents. *Nature, 413,* 211–218.

Mombaerts, P. (2004). Genes and ligands for odorant, vomeronasal and taste receptors. *Nature Reviews. Neuroscience, 5,* 263–278.

Munger, S. D., Leinders-Zufall, T., & Zufall, F. (in press). Subsystem organization in mammalian olfaction. *Annual Review of Physiology, 71,* 3.1–3.26. (Epub ahead of print) Online DOI: 10.1146/annurev.physi01.70.113006.100608

OLFACTORY STIMULUS

In order for us to perceive something, there must be a source that triggers our perception. That source is called a *stimulus*. A stimulus is composed of some kind of energy that our senses are tuned to detect. For example, the visual system can detect electromagnetic energy, that is, light, and the auditory system can detect vibrational energy, that is, sounds. In turn, the sense of smell, called olfaction, is tuned to detect chemical energy stored in some particular compounds called odorants. The odorants constitute the olfactory stimuli. This entry covers units of change, management of the olfactory stimulus, complexity of the olfactory stimulus, and sniffing.

Units of Change of the Olfactory Stimulus

The discovery of the gene family of olfactory receptors in 1991 led to an accelerated and much better understanding of the olfactory sense. Among other issues, this progress clarified the following: (a) the series of molecular events that take place within olfactory sensory neurons after the odorants bind to the olfactory receptors, and (b) the organization of the anatomical connections through which the information is transmitted from the periphery to the olfactory bulb and higher levels of the olfactory system to produce an odor perception. The data supported the notion that, typically, each olfactory neuron uses one particular type of receptors, and each receptor type responds to more than one odorant. This led to the conclusion that even a single kind of odorant molecule produces a neural message composed of the combined output (pattern) of many receptors with overlapping odorant specificity.

Despite these advances, little is known about the details of the interaction between an odorant and its receptors or about the specific structure or molecular properties that make an odorant a strong olfactory stimulus (i.e., one with a low odor detection threshold) or a weak olfactory stimulus (i.e., one with a high odor detection threshold). For humans, and for terrestrial animals in general, odorants need to have, first of all, enough volatility to reach the olfactory receptors in the nose, because they enter as vapors carried by the air breathed. As discussed in a later section, *sniffing,* the act of increasing the airflow through the nasal passages, is an important component of the olfactory process.

A problem encountered in trying to establish structure-activity relationships between odorants and olfactory outcomes, such as odor detection thresholds, odor discrimination, and odor quality, is the lack of understanding of the unit(s) of chemical change that underlies the olfactory response. In olfaction, we lack the equivalent of the wavelength range of light in vision or the vibrational range of sounds in hearing. One strategy employed to ameliorate this deficiency has been to test the olfactory outcome (e.g., threshold, discrimination) across homologous series of chemicals and across chemical functionalities. In this way, the number of carbon atoms in the chain, in the first approach, and the particular chemical functional group (e.g., alcohol, ester, ketone), in the second approach, become a parameter that partially reflects units of chemical change.

Studies of structure-activity relationships in olfaction have been more successful when limited to a particular type of odor (i.e., ambergris or musk) than when attempting to provide an explanation across the broad spectrum of odor qualities. There are odorants that, despite having very different chemical structures, still possess similar odors. For example, hydrogen cyanide and benzaldehyde are radically different molecules, they both smell like almonds to humans. Conversely, there are odorants very similar in structure that evoke dissimilar odors. For example, *d*-carvone and 1-carvone are practically identical molecules except that they cannot be superimposed because they are mirror images of each other (called *enantiomers* in chemistry). To illustrate, the effect is similar to the way in which the left and right hands are also mirror images of each other and cannot be superimposed. Still, *d*-carvone smells like peppermint and 1-carvone smells like caraway, two similar but discriminable odors to humans. The bases for these contrasting effects regarding chemical structure similarity and odor quality are not completely understood.

Other investigations have focused on describing the physicochemical basis for odor detection, that is, odorant potency measured as thresholds, rather than for odor quality. This seems a productive initial approach because detection entails a simpler neural phenomenon than identification and probably rests on simpler general principles. In any case, whatever structure-odor relationships are deduced, there always seem to be exceptions. It is quite likely that the study and understanding of such exceptions will lead to a better knowledge of the topic.

Managing the Olfactory Stimulus

The lack of standardized equipment to handle the olfactory stimulus (i.e., the odorant vapor) and its inherent nature make its generation, control, delivery, and quantification more problematic than, for example, the visual stimulus (i.e., light) or the auditory stimulus (i.e., sound). The techniques used to handle and measure olfactory stimuli are called *olfactometry,* and the devices employed are called *olfactometers.* Some olfactory studies resort to nominal or indirect quantification of odorant vapors. For example, they might only measure the concentration of the liquid phase containing the odorant, but not that of the vapor phase that constitutes the actual stimulus. Thus, the vapor is either not quantified, or its concentration is theoretically calculated from values of vapor pressure that can vary widely among data sources, and that might be themselves the result of calculations. In other cases, quantification of the odorants takes the form of percentages of dilution with air from a source that is assumed to have a certain concentration (e.g., vapor saturation at room temperature: ≈23 °C), but that is not experimentally measured, for example, by gas chromatography. Although the nature and scope of some investigations might not require an absolute and stringent chemical quantification of the odorant(s), progress in understanding many aspects of olfactory function rests on an accurate experimental measurement of the actual olfactory stimulus, that is, the concentration of the odorant vapor. Although the human nose is often more sensitive and sophisticated than many apparatus for chemical measurement, the application of existing and emerging powerful chemical-analytical instruments to the management of the olfactory stimulus needs to be more common and widespread.

To achieve optimum olfactometric measurements, the quantification of the olfactory stimulus needs to be complemented with an appropriate availability of the stimulus to the tested subjects. It has been reported that an "average" human sniff has an inhalation rate of 30 liters per minute (1/min), a volume of 200 milliliters (ml), and a duration of 0.4 seconds (s). If the olfactometer and technique employed for stimulus delivery does not meet the input required by the subject, the measurements obtained might not have full ecological relevance.

Complexity of the Olfactory Stimulus

Many studies of the sense of smell use a single type of odorant molecules as olfactory stimuli. The variability in structure and chemical functionality of individual odorants is enormous. Many odorants could be classified as volatile organic compounds, but quite a few of them are semivolatile (the border between volatile and semivolatile is lax and fuzzy), and still others are small inorganic molecules (e.g., ammonia, hydrogen sulfide). Some investigations have used mixtures of two or more odorants to address the issue of how olfaction process mixtures of odorants. In environmentally realistic situations, the olfactory stimulus is composed of dozens or even hundreds of individual odorants that are often perceived as a unity, for example, the odor of coffee, chocolate, soap, tobacco, and rubber.

The rules of how mixtures of odorants are perceived as regards to odor detection, odor discrimination, and odor quality have not been clarified yet. Nevertheless, some general principles have been suggested. At low levels of detection, individual odorants tend to add their individual detectabilities to generate the overall detection of the mixture to a larger degree than they do at higher levels of individual detectability. Regarding odor discrimination, a number of studies found that humans could identify a maximum of four individual components in odor mixtures. This result did not seem to be altered with the training and experience of the subjects, the type of odorants mixed (i.e., good or bad "blenders"), or whether the individual "odors" were single chemicals or complex mixtures (e.g., cheese, honey). Regarding odor quality, mixtures of odorants can be perceived as different from the individual components or, in other cases, the components can retain their individual quality when mixed. Between these two extremes, a whole array of intermediate outcomes is also possible. Although the rules for the production of a particular result are not completely understood, the outcome has been shown to depend on chemical and perceptual similarity among components, on their relative concentration in the mixture, and on olfactory receptor overlap.

The Sniff and the Olfactory Stimulus

Studies on the effect of sniff parameters on various olfactory outcomes indicated that the sniff plays an important role in odor detection, intensity, and identification. The parameters include sniff flow rate, volume, duration, interval, and number. The issue is important for olfactometry because any effort to standardize odorant delivery should make sure that instruments and methods provide the appropriate conditions for the particular olfactory task in the specific experimental context. Not surprisingly, natural sniffing produces optimum odor perception. It has been suggested that (a) sniffing influences olfactory neural activity in the olfactory bulb and cortex, (b) there is a temporal synchrony between sniffing and the actual odorant-induced neural activity along various levels of the olfactory pathway, and (c) there is a dedicated olfactomotor system, although little is known about the neural mechanisms that control it.

The importance of sniffing in olfaction rests in that the sniff determines the spatial and temporal distribution of the olfactory stimulus over the olfactory epithelium. Reciprocally, the olfactory stimulus has also been shown to alter sniffing behavior. For example, malodors produce a reflex-like reduction in sniffing, and stronger intensities of an odor induce lesser-volume sniffs. Within limits, odorant detection increases with sniff flow rate but does not change with sniff volume as long as a minimum 200 ml is sniffed. It has been suggested that the perceived intensity of a fixed odorant concentration remains constant despite varying sniff flow rate because the olfactory system produces a "correction" that accounts for perceived effort during the sniff. This model, however, might not generalize to all odorants or conditions.

Sniff flow rate can influence odor quality because, as odorant vapors travel above the olfactory mucosa, their deposition pattern will differ according to their tendency to dissolve, or sorb, in the mucosa. Low-sorption odorants will distribute uniformly across the mucosa at low rates, but unevenly (accumulating posteriorly) at high rates. In contrast, high-sorption odorants will distribute unevenly (accumulating anteriorly) at low rates, but uniformly at high rates.

J. Enrique Cometto-Muñiz

Acknowledgments: Preparation of this entry was supported by grant number R01 DC002741 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

See also Olfaction; Olfactometry; Olfactory Adaptation; Olfactory Receptors and Transduction; Perceptual Development: Taste and Olfaction

Further Readings

Cometto-Muñiz, J. E. (2001). Physicochemical basis for odor and irritation potency of VOCs. In J. D. Spengler, J. Samet, & J. F. McCarthy (Eds.), *Indoor air quality handbook* (pp. 20.21–20.21). New York: McGraw-Hill.

Cometto-Muñiz, J. E., & Abraham, M. H. (2008). Human olfactory detection of homologous n-alcohols measured via concentration-response functions. *Pharmacology, Biochemistry and Behavior, 89,* 279–291.

Cometto-Muñiz, J. E., Cain, W. S., & Abraham, M. H. (2003). Quantification of chemical vapors in chemosensory research. *Chemical Senses, 28,* 467–477.

Goyert, H. F., Frank, M. E., Gent, J. F., & Hettinger, T. P. (2007). Characteristic component odors emerge from mixtures after selective adaptation. *Brain Research Bulletin, 72,* 1–9.

Laing, D. G. (1986). Identification of single dissimilar odors is achieved by humans with a single sniff. *Physiology and Behavior, 37,* 163–170.

Mainland, J., & Sobel, N. (2006). The sniff is part of the olfactory percept. *Chemical Senses, 31,* 181–196.

Malnic, B., Hirono, J., Sato, T., & Buck, L. B. (1999). Combinatorial receptor codes for odors. *Cell, 96,* 713–723.

Wise, P. M., Miyazawa, T., Gallagher, M., & Preti, G. (2007). Human odor detection of homologous carboxylic acids and their binary mixtures. *Chemical Senses, 32,* 475–482.

Zarzo, M. (2007). The sense of smell: Molecular basis of odorant recognition. *Biological Reviews of the Cambridge Philosophical Society, 82,* 455–479.

OPTIC ATAXIA

In normal adults, even the most mundane interactions with the objects around us are virtuoso feats of elegance and efficiency. Much of this elegance derives from the effortless use of vision to anticipate the properties of objects and to monitor the progress of ongoing movements. In reaching for a glass of water, for instance, visual localization helps you direct your hand to the glass, and visual analyses of its size and shape guide your hand to enclose it. Vision may also prompt you to veer around other objects or to correct your movement for errors or for changes in the position of the glass or obstacles. *Optic ataxia* describes a condition in which some or all of these abilities are lost following damage to the parts of the brain linking vision to action. It is broadly characterized by an inaccuracy of visually guided movement, which is not due to more general visual or motor impairments. In reaching for a seen object, a person with severe optic ataxia may seem to grope in the dark, extending a flattened hand hesitantly until a chance contact allows the object to be retrieved by touch. This entry focuses on the history, presentation and diagnosis, additional features, and neuroanatomy of optic ataxia

History

The term optic ataxia (*optische Ataxie* in the original German) was coined by the Hungarian physician Rudolf Bálint in his 1909 report of a man with lesions of the posterior parietal lobe on both sides of the brain. Optic ataxia was one of a cluster of symptoms, now known collectively as Bálint's syndrome, which included the restriction of visual attention to single objects and a paucity of spontaneous eye movements. Bálint noted inaccurate reaching of the right hand, and the patient himself commented that this hand was clumsy. For instance, he often found himself lighting a cigarette at its middle instead of at its end. These errors could not be ascribed to any general visual impairment, as his left-hand reaching was accurate. Similarly, he could make coordinated right-hand reaches, with eyes closed, to different parts of his body, ruling out a right-sided movement disorder. Thus, the problem was neither visual nor motor, but *visuomotor,* emerging only for movements made under visual guidance. This confluence is well captured by Bálint's coinage, which implies a movement disorder (ataxia) that is specifically visual (optic).

Presentation and Diagnosis

When optic ataxia is severe, misreaching will be obvious to the patient and to others. However, specialized testing may be required to confirm the symptom or to diagnose its subtle forms. Typically, the examiner will present an object, such as a pen, to the left or the right side for grasping by each hand. This will be done both when the patient can look at the pen directly and when they must look straight ahead so that it falls outside of central vision. Misreaching most often emerges only in the latter case, with accurate reaching under direct viewing. It may also be specific to certain combinations of the target side and hand. Two main patterns can be expected: a *field effect,* whereby large errors are made, with either hand, on the side of space opposite the responsible brain damage; and/or a *hand effect,* whereby the hand anatomically opposite to the brain damage misreaches for targets on either side. To confirm the diagnosis, it should be shown that the errors are not due to other disorders affecting vision or movement.

Additional Features

The sine qua non of optic ataxia is inaccurate reaching to visual targets, but a range of related impairments may be present, including failure to rotate or preshape the hand appropriately for grasping, inability to correct ongoing movements for changes in target position, and insensitivity to obstacles in the reach space. Inaccurate visuomotor guidance may also be observed for the foot by asking the patient to touch a target with the toe. No systematic study has investigated whether these subsymptoms require damage to different brain areas. For misreaching, at least, the pattern of errors may depend on the side of the brain damage. Marie-Therese Perenin and A. Vighetto noted that damage to the right side of the brain tended to provoke a field effect, while an additional hand effect was often present after left-brain damage. However, because the number of patients studied was small, the generality of these patterns awaits confirmation. The typical spatial character of the errors is also uncertain: Several investigators have found that, for peripheral targets, optic ataxic misreaching tends to be markedly toward central (looking straight ahead) vision, though not all patients have been found to show this pattern.

Neuroanatomy

Patients with optic ataxia may have extensive brain damage, especially if the symptom manifests as part of Bálint's syndrome. Bálint's original patient, for instance, had lesions to the posterior parietal lobe, extending into the occipital lobe, on both sides of the brain. However, lesions restricted to one side of the brain are the more usual cause of isolated optic ataxia. By studying patterns across patients, it may be possible to discern whether there is any specific subregion that is invariably affected. The key areas of involvement have usually been found to be the intraparietal sulcus and adjacent parts of the superior parietal lobe. However, the latest and largest lesion analysis has suggested that misreaching for peripheral targets may more often follow damage close to the junction between the occipital and parietal lobes. These brain areas are presumably critical for the visual guidance of reaching in humans.

Robert D. McIntosh

See also Action and Vision; Eye and Limb Tracking; Neuropsychology of Perception; Reaching and Grasping

Further Readings

Karnath, H. O., & Perenin, M. T. (2005). Cortical control of visually-guided reaching: Evidence from patients with optic ataxia. *Cerebral Cortex, 15,* 1561–1569.

Perenin, M. T., & Vighetto, A. (1988). Optic ataxia: A specific disruption in visuomotor mechanisms. *Brain, 111,* 643–674.

Pisella, L., Gréa, H., Tilikete, C., Vighetto, A., Desmurget, M., Rode, G., et al. (2000). An "automatic pilot" for the hand in human posterior parietal cortex: Toward reinterpreting optic ataxia. *Nature Neuroscience, 3,* 729–736.

Schindler, I., Rice, N., McIntosh, R. D., Rossetti, Y., Milner, A. D. (2004). Automatic avoidance of obstacles is a dorsal stream function: Evidence from optic ataxia. *Nature Neuroscience, 7,* 779–784.

OSCILLATORY SYNCHRONY

From a neural point of view, building a coherent percept can be compared to a jigsaw puzzle. The pieces of the puzzle are the bits of sensory information that are scattered in different sensory areas: the color of an object, for instance, is analyzed in one brain region, and its motion in another one. But how are these pieces put back together again to form the coherent percept we experience? This problem, also known as the feature binding problem, could be solved by *oscillatory synchrony:* All neurons dealing with the same perceptual object would synchronize their firing on an oscillatory mode—in other words, alternate collectively between epochs of activity and rest—much as a group of drummers would follow a given rhythm. In this view, the temporal structure of neural activity becomes a fundamental property of the neural code.

Binding features together in an object is crucial, but distinguishing objects from one another is equally important. Oscillatory synchrony may be relevant for scene segmentation: In response to a picture of a woman carrying a dog, for instance, neurons responding to the woman would synchronize their discharge at a given frequency, whereas neurons responding to the dog would synchronize their discharge at a different frequency—as two groups of drummers could beat two distinct rhythms. In addition to scene analysis, oscillatory synchrony may also prove useful for other cognitive functions, such as attention, memory, and awareness. This entry describes neural oscillations and measuring oscillatory potentials.

Neural Oscillations

A typical experimental test of the role of oscillatory synchrony in feature binding involves two physically similar perceptual configurations: one leading to a coherent percept (e.g., dots defining the contours of a triangle) but not the other (randomly placed dots). Oscillatory synchrony is larger in response to the coherent stimulus: It could thus be involved in grouping together different sensory features in a coherent perceptual whole. This finding has been replicated a number of times in the visual modality, using a variety of stimulus configurations and recording methods, in both humans and animals. The role of oscillatory synchrony in other sensory modalities, as well as in cross-modal integration (e.g., integrating the sound of a bark with the picture of a dog), is also well documented. Direct tests of the role of oscillatory synchrony in segmentation are more scarce, especially in humans.

Because oscillatory synchrony has been primarily involved in feature binding, it has been proposed as a mechanism linking direct neural activity to subjective experience. Thus, it has often been considered as a potential neural correlate of (or prerequisite to) awareness. Oscillatory synchrony shows an additional interesting feature: It is by definition a group property, in line with the idea that awareness cannot be reduced to activity in a single module, area, or neuron, but emerges from dynamical interactions within the brain. This appealing theoretical proposal remains quite difficult to validate experimentally in an irrefutable manner, although a growing body of evidence points in this direction.

The oscillations related to grouping and awareness are quite fast, in what is called the gamma range, between 30 and 100 hertz (Hz). This means that an oscillation period lasts between 10 and 30 milliseconds (ms). This corresponds to a temporal window of crucial importance for neural integration: All inputs arriving on a given neuron within 10 to 30 ms add up and are likely to generate a response in this neuron. On the other hand, the same number of inputs arriving over a longer period of time (e.g., 100 ms) will not add up to generate a response. In other words, inputs that are synchronized are more likely to be selected for further processing than nonsynchronized inputs: This is the definition of an attentional selection mechanism. Experimentally, attention-related gamma oscillations have indeed been described, both in monkeys and humans. In addition, synaptic plasticity—the fundamental mechanism by which learning takes place and new memories are formed—can be highly sensitive to input timing, in particular within 10 to 30 ms. A growing number of experimental data in humans and monkeys corroborate the involvement of oscillatory synchrony in memory formation, in both the visual and auditory modality. For instance, items that elicit larger gamma oscillations when seen for the first time are more likely to be recalled later.

Measuring Oscillatory Potentials

How is oscillatory synchrony measured? When a sufficiently large number of neurons get synchronized, their electrical activity can be picked up in population recordings. Typically in these recordings, oscillatory synchrony appears as a sine wave, with a regular alternation of peaks and troughs reflecting successive states of collective excitation and inhibition.

Oscillations in the gamma range have been observed in small (~10) to medium (~100 to 1000) and large (~10,000) groups of neurons, in all sensory modalities, in many different species, from locusts and turtles to monkeys and humans. Both the existence and functional role of oscillatory potentials are nevertheless regularly challenged. There are two main reasons for this skepticism. The first one is technical: Oscillatory potentials remain quite difficult to record and analyze, in particular in the gamma range because they are small and can be mistaken with various types of artifacts. The second reason is more conceptual: Oscillatory potentials pertain to the neural code (i.e., the very nature of information in the brain). The challenging theoretical hypothesis that time is an essential feature of the neural code is not acknowledged by everyone yet.

A Large Spectrum of Oscillations

Oscillations are certainly not confined to the gamma range, although this frequency range has attracted much attention recently. Historically, the low-frequency, large amplitude oscillations were the first to be described. It would be beyond the scope of this entry to review extensively those oscillations, but the potential relevance of slower rhythms for perception should not be forgotten. For instance, oscillations in the alpha range (between 8 and 12 Hz) that naturally stop when opening the eyes are often considered as a marker of sensory reactivity, and numerous results point toward a significant influence of alpha rhythms on perception. Current research aims at integrating the different time scales of perception and the different frequencies of oscillatory potentials in a comprehensive schema.

Catherine Tallon-Baudry

See also Attention: Physiological; Attention and Consciousness; Binding Problem; Consciousness; Neural Representation/Coding

Further Readings

Buzsaki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science, 304,* 1926–1929.

Jensen, O., Kaiser, J., & Lachaux, J .P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends in Neuroscience, 30,* 317–324.

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neurosciences, 18, 555–586.*

Tallon-Baudry, C. (2009). The roles of gamma-band oscillatory synchrony in human visual cognition. *Frontiers in Bioscience, 14*(1), 321–332.

OTOACOUSTIC EMISSIONS

In addition to receiving and processing sound stimuli, the ear is capable of producing sounds too. First described by British physicist David Kemp in 1978, *otoacoustic emissions* (OAEs) are low-level sounds that are generated within the cochlea of the inner ear. These weak vibrations are transmitted outwardly through the middle ear to the ear canal, where they can be recorded with a sensitive microphone. This entry describes the mechanism that creates OAEs and the clinical use of OAEs.

Otoacoustic emissions arise from the miniscule movements of the outer hair cells (OHCs) in the cochlea. The OHCs structurally resemble the inner hair cells, which are true sensory cells in that they convert vibrations into neuroelectrical responses the brain understands. The OHCs, on the other hand, act as tiny mechanical amplifiers. These cells are attached to the basilar membrane, which is a flexible ribbon of tissue that extends the entire length of the cochlea. When the ear is stimulated with sound, the basilar membrane undulates—moving up and down several thousand times per second. The OHCs, in turn, react by changing their length extremely rapidly, becoming shorter and longer via contraction and expansion, respectively. The push-pull action exerted by the OHCs boosts the basilar membrane's motion, and this enhanced vibration results in improved hearing sensitivity.

The typical paradigm for evoking OAEs involves presenting a series of brief acoustic clicks or a steady-state tone to the ear and then recording the ensuing sound. The click method produces transient evoked otoacoustic emissions (TEOAEs). Because the stimulus contains many frequencies, a large number of frequencies are present in TEOAE recordings. The tonal technique yields stimulus-frequency otoacoustic emissions (SFOAEs). SFOAEs are not in widespread use, as it is difficult to separate the response generated in the ear from the stimulus being presented. A more common way to elicit OAEs is by delivering not one but two tones simultaneously. The tones generate patterns of vibration that interact on the basilar membrane. Through this mixing process, the wave patterns are altered slightly and so-called distortion-product otoacoustic emissions (DPOAEs) are produced. Alternatively, spontaneous otoacoustic emissions (SOAEs) occur in the absence of external stimulation. SOAEs occur because the OHCs in a restricted region on the basilar membrane appear to vibrate continuously without provocation. Localized areas of the basilar membrane correspond to specific pitches, and thus SOAEs are nearly tonal in character.

As mentioned previously, the lengthwise motion of the OHCs plays a key role in generating all types of OAEs. At a more fundamental level, at least two different mechanisms may contribute to the production of OAEs. One underlying mechanism involves mechanical reflection. Slight irregularities are thought to exist at various locations along the basilar membrane. Perhaps, these imperfections serve as acoustic reflectors and inbound sound simply bounces off these microscopic flaws in essentially the same way light reflects off a mirror. The second mechanism is based on nonlinear distortion. The amplification provided by the OHCs depends on the level of the incoming sound—soft sounds are boosted more than loud sounds. One consequence of this compressive phenomenon is that stimulus vibrations become "distorted" such that frequencies not present in the original sound are introduced into the response. The best way to observe cochlear distortion is to deliver two tones and then look for more than two frequency components in the spectrum of otoacoustic response. Moreover, the mechanisms of reflection and distortion likely both contribute to the production of evoked OAEs. Work is currently underway to determine the relative contribution of each mechanism.

Evoked OAEs are emitted by ears with a full complement of functional OHCs. If these cells are damaged by disease or trauma or rendered ineffective through anomalous genetics, OAEs are not produced. This relationship makes OAE measurement a valuable clinical procedure for evaluating cochlear health. The recording procedure is simple, fast, and noninvasive. These features are especially important when assessing individuals who are either unable (e.g., infants or people with motor impairment) or unwilling to be truthful (e.g., people who stand to benefit from the diagnosis of hearing loss).

Because most cases of permanent hearing loss have a cochlear site of lesion, evoked OAEs are widely used for hearing screening of newborns. Additionally, certain medications have toxic side effects on OHCs. Evoked OAEs are used to monitor changes in OHC function during drug treatment so that attempts to minimize drug-related hearing loss can be undertaken. Recording OAEs does not require direct participation by the patient, and the patient cannot control the OAE response. Because OAE presence suggests vigorous OHCs, one can rule out feigned hearing loss in cases where OAEs are observed. One final caveat regarding clinical measures of OAEs deserves mention. Although OAE presence suggests normal cochlear function, this outcome does not necessarily imply normal hearing. It is certainly possible that damage to the neural pathways going to the brain could exist independent of cochlear insult. This could lead to compromised auditory function and perhaps hearing loss, even though the inner ear is undamaged and normal OAEs are observed.

Craig A. Champlin

See also Audition: Disorders; Auditory Processing: Peripheral; Auditory Receptors and Transduction; Physiological Approach

Further Readings

Brownell, W. E. (1990). Outer hair cell electromotility and otoacoustic emissions. *Ear and Hearing, 11,* 82–92.

Kemp, D. T. (1978). Stimulated acoustic emissions from within the human auditory system. *Journal of the Acoustical Society of America, 64,* 1386–1391.

Shera, C. A. (2004). Mechanisms of mammalian otoacoustic emission and their implications for clinical utility of otoacoustic emissions. *Ear and Hearing, 25,* 86–97.

OUT-OF-BODY EXPERIENCE

An *out-of-body experience* (OBE) is a transient impression of being separated from one's own body. Usually it is associated with three main phenomenological elements: the impression of the self being localized outside one's body (disembodiment), the impression of seeing the world from this elevated perspective (altered first-person perspective), and seeing one's own body below (autoscopy). This entry describes OBE's observed in clinical settings and how OBE has been studied experimentally in healthy participants.

These are, of course, unusual and uncommon experiences, because we normally perceive our self as localized within our bodily borders (embodiment) at a distinct location in space from where we also perceive the world (first-person perspective). Therefore, the scientific study of OBEs offers a unique possibility to gain insights into the functional and neural mechanisms of embodiment, the first-person perspective, and related aspects of self-representation.

OBEs occur in a variety of conditions, such as during sensory deprivation, drug use, certain sleep states, as well as in some forms of psychiatric and neurological disease. OBEs may also occur in healthy people. In particular, the study of neurological OBEs due to epilepsy, migraine, or vascular stroke has revealed some of their functional and neural mechanisms. OBEs of neurological origin are part of a group of disorders that are characterized by illusory reduplication of one's body, but with various degrees of illusory self-location that are associated with vestibular dysfunction. OBEs have been associated with brain damage in the temporoparietal cortex, an area that is strongly involved in multisensory integration of bodily information, such as visual, tactile, and vestibular information about one's own body. Intracranial stimulation of the temporoparietal cortex was shown to evoke vestibular sensations, visuo-tactile, visuo-proprioceptive, or kinesthetic illusions, as well as full-blown OBEs in epileptic patients. These data have highlighted the crucial role of multisensory bodily representations in the temporoparietal cortex in OBEs and self-representation.

Clinical studies do, however, present certain experimental limitations, such as small sample size and limited generalization. Furthermore, they do not explain how the experience of a unitary and stable self (including embodiment and first-person perspective) is achieved under normal conditions. Therefore, several recent studies have utilized knowledge about the mechanisms and phenomenology of OBEs to investigate the influence of multisensory conflicts on bodily self-representations in healthy participants. In 1899, G. M. Stratton carried out self-experiments walking himself with a portable device of mirrors that were aligned in such a way that he could see a projection of his body below him, creating a set-up that allowed him to induce multisensory and sensorimotor conflicts between the seen and the felt position of his body. After wearing this device for ~24 hours, he reported that he "was mentally outside his own body." He argued that the visual information about his body dominated the other bodily information (also known as "visual capture").

In two recent experiments, the influence of multisensory conflicts on global self representations was tested more systematically using a video camera and head-mounted displays (HMDs). In one study, the sitting participant was stroked on the chest with a stick and filmed from the back. On the HMD, participants could see the stick moving in front of the camera, and this generated the feeling of being located at the position of the camera (outside the own body), which was also supported by psychophysiological measurements. In a similar study, standing participants were filmed from behind and saw—via the HMD—either their own body, a fake body, or an object being stroked synchronously or asynchronously with their own body. In conditions when the stroking was synchronous and applied to a bodily (fake or own) character, participants self-identified with the body in front of them and localized the self as drifted toward the front (and thus to a position outside the bodily borders). This suggests that visual capture may influence global body representations, leading to disturbed self-identification, disturbed spatial unity, and, under some conditions, the first-person perspective.

Even though none of these experimental studies induced full-blown OBEs, the results support the conclusions drawn from clinical data obtained in neurological patients in suggesting that OBEs and related phenomena might occur when the brain is exposed to conflicting multisensory information from the different bodily sensory inputs and is no longer able to assemble them into a coherent global body representation. This results when brain areas, such as the temporoparietal cortex, that are involved in this integration process under normal conditions, are damaged or when conflicting information is given under unusual, yet realistic, experimental setups. Virtual reality (VR) may be the most efficient method to generate such experimental setups, and previous researchers have pointed out commonalities between OBEs, VR, and presence (defined as the feeling of being located in a different [virtual] place and behaving accordingly). Howard Rheingold, for example, mentions explicitly that presence is similar to OBEs. Furthermore, VR is a useful tool for studying the functional mechanisms of OBEs because it enables the systematic and yet realistic manipulation of the first-person perspective, global own body representations, and the congruency between multisensory bodily cues. At present, studies investigating global bodily self-consciousness in healthy participants and neurological patients are sparse. The merging of techniques from experimental psychology and VR with the wide array of cognitive neuroscience methods is likely to further the understanding of how the temporoparietal cortex and other brain regions are involved in the generation of OBEs, related phenomena, and the first-person perspective of the embodied conscious self.

Olaf Blanke and Bigna Lenggenhager

See also Body Perception; Body Perception: Disorders; Consciousness, Consciousness: Disorders; Hallucinations and Altered Perceptions; Pain: Neuromatrix Theory; Phantom Limb

Further Readings

Blackmore, S. J. (1982). *Beyond the body. An investigation of out-of-body experiences.* London: Heinemann.

Blanke, O., Landis, T., Spinelli, L., & Seeck, M. (2004). Out-of-body experience and autoscopy of neurological origin. *Brain, 127,* 243–258.

Ehrsson, H. H. (2007). The experimental induction of out-of-body experiences. *Science, 317,* 1048.

Irwin, H. J. (1985). *Flight of mind: A psychological study of the out-of-body experience.* Metuchen, NJ: Scarecrow Press.

Lenggenhager, B., Tadi, T., Metzinger, T., & Blanke, O. (2007). Video ergo sum. Manipulating bodily self-consciousness. *Science, 317,* 1096–1099.

P

PAIN: ASSESSMENT AND MEASUREMENT

Pain in the clinic cannot be physically observed and usually cannot be independently precisely controlled. Such control is possible in laboratory studies in which pain is induced by controlled stimulation. This entry discusses how pain is assessed and measured in the laboratory as well as in the clinic.

Measurement of Pain in the Laboratory

Experimental heat, mechanical, electrical, chemical, and other forms of stimulation are extremely useful in studies of pain in both man and animals. Pain is inferred in animals from a range of physiological measurements and from reactive and operant behaviors. Those uncomfortable with human verbal reports of pain have promoted the same type of measurements for pain in humans using physiological measurements (such as spinal reflexes, cortical evoked potentials, or functional neuroimaging) and behavioral indices ranging from observed measures of grimacing and bracing to measures of social and work activities. However, pain is only truly defined in humans, based on the unique ability to describe and evaluate private experience.

Laboratory methods provide a degree of control over the subjective measurement process that is not possible with clinical pain measurement. These methods demonstrate that human subjects can reliably rate pain, distinguish between dimensions of sensory intensity and unpleasantness, and discriminate different pain qualities. They also provide valuable information about rating scale behavior that can be applied to clinical pain assessment. Many traditional psychophysical procedures, which focus on relations between stimulus-evoked sensations, have been applied to the evaluation of pain sensation. These methods can be divided into those that assess the pain threshold and those that evaluate the suprathreshold range from pain threshold to pain tolerance.

The term *pain threshold* is often used to describe general pain sensitivity, such as “he has a high pain threshold.” The specific meaning of pain threshold refers to the distinction between sensory qualities of nonpainful and painful stimulation. As the magnitude of a potentially painful stimulus is increased, the pain threshold marks the transition from the absence of pain sensation to the presence of pain sensation, and is quantified as the amount of stimulus intensity needed to evoke a painful sensation. In the laboratory, the pain threshold is determined by established psychophysical procedures that minimize biases that influence the result. These include the classical method of limits and the method of constant stimuli. The method of limits uses alternating ascending and descending series of continuous or discrete stimuli. These alternating series control for directional errors such as anticipation, of indicating the presence of pain before the sensation becomes painful, and errors of habituation, of continuing to use the nonpainful responses after the stimulus has become painful. To further control biases, both the starting point and amount of change in successive stimuli can be randomly varied. For assessment of the pain threshold, this method is usually modified to use only ascending series to prevent unacceptably painful stimulation at the beginning of a descending series. The method of constant stimuli presents discrete stimuli of varying intensities that span the range from certain judgments of nonpain to certain judgments of pain. Thus, these stimulus values span the "gray area" of uncertain responding. The result of this method is usually plotted as a probability function extending for 0 to 100% probability, and a stimulus value corresponding to a specific value (such as 50 or 75%) is defined as the pain threshold.

Threshold methods are limited in that they only assess sensitivity to weak pain sensations at the very bottom of the perceptual pain range. This sensitivity may not reflect sensitivity throughout the range from pain threshold tolerance, which can be assessed by methods that evaluate subjective judgments of suprathreshold pain sensations.

Suprathreshold methods deliver sensations over the entire perceptual range from threshold to tolerance. Most of these deliver a series of discrete stimuli that vary randomly in intensity sequence. The dependent variable is the response made using a subjective rating scale. The difference in the varying stimulus intensities should be small enough to create confusion between adjacent stimulus intensities, forcing the subject to make ratings based on judgments of sensation and not on identification of a specific stimulus (e.g., that is the middle stimulus that I call "5"). These procedures vary in the type of response scale and the type of analysis. The response scales range from the ubiquitous numerical 1 to 10 scale and simple category scales (mild, moderate, severe) to continuous measures such as the visual analog scale (VAS) that involves marking a line that represents a range of pain sensation. The most common example of a VAS is a 10-centimeter (cm) horizontal line labeled at the extremes by "no pain" and "most intense pain imaginable" or some other descriptor of intense pain. Another approach combines these methods by spacing quantified descriptors along a continuous scale or numerical category scale. These scales have traditionally been presented on paper but modern versions are delivered by computer or handheld PDA (personal digital assistant) devices. All of these methods have been shown to be reliable and valid by repeated use in pain studies. One of the issues in the selection of a method, especially with patients in pain, is the selection of the appropriate stimulus values. In addition, the use of a fixed stimulus set to evaluate analgesics provides an undesirable cue because all of the stimuli will be less intense. These issues have been addressed by another class of suprathreshold scaling methods that use interactive computer algorithms to continuously adjust varying stimulus intensities to produce the same level of subjective sensations in all subjects. These methods can be considered to be a second-generation approach because of their interactive nature. These methods automatically adjust the stimulus levels to each participant, avoiding problems of selecting appropriate stimulus levels for each person. They also provide the same subjective levels of pain before and after an analgesic intervention, minimizing extraneous clues that an active drug has been administered.

Further issues involved in selection of a suprathreshold scaling method include applicability to multisite international studies (a problem with verbal scales), ability to provide absolute judgments of pain magnitude (greater face validity with verbal scales), and the need for a motor response (line marking methods are problematic for telephone evaluations and for studies such as brain imaging in which a motor response is difficult or undesirable). Responses that are easily remembered, such as simple category scales, may not be the best choice for measures of repeat reliability, because it is not clear if the subject remembers the pain or the previous response. Another issue concerns the ability to rate significant pain dimensions.

To assess the relevant dimensions of pain, measurement systems must go beyond the assessment of sensation. Pain is not only a sensation but a negative feeling state that motivates behavior. This is a problem also for the assessment of other sensations with homeostatic consequences, such as temperature, taste, and olfaction that contain both negative and positive feeling states. Measurement methods used for these homeostatic sensations have been adapted to the assessment of pain, distinguishing between the distinct dimensions of pain intensity and pain unpleasantness, and have shown how these dimensions differ among different types of pain and respond differently to pain control interventions. For example, there is evidence that verbal scales with words descriptive of a single pain dimension (e.g., mild, moderate, intense for the dimension of pain intensity and uncomfortable, annoying, distressing for pain unpleasantness) facilitate the discrimination of that dimension. The types of studies demonstrating these properties have shown that an intravenous opioid reduces pain intensity with variable effects on unpleasantness and that the amount of unpleasantness associated with a specific level of sensory intensity is less in experimental situations and more in chronic pain conditions, which show further differences between conditions such as fibromyalgia and cancer.

Pain psychophysical methods are often criticized because the laboratory environment does not duplicate the sensory and emotional aspects of the clinical situation. While true to some extent, these procedures have been extremely useful for increasing the knowledge about the neural mechanisms that mediate pain and analgesia. The application of these methods to patients, termed quantitative sensory testing (QST), essentially extends the neurological examination to discover abnormalities in simple pain perception and in processing of prolonged pain. QST has greatly increased the capabilities of clinical pain assessment, leading to both improved diagnosis and treatment and to improved knowledge of poorly understood pain mechanisms. QST involves the same methods as used in the laboratory, with the complexity of these methods matched to the alteration of symptoms. Gross changes can be detected by simple procedures, such as raising the temperature of a probe slowly until a report of pain is made. More subtle symptoms may require the more elaborate procedures previously described. The modalities commonly used are thermal (both hot and cold), pressure (both blunt and punctuate), and vibration. Other methods include nonpainful and painful electrical stimulation and methods such as repeated heat stimuli that activate spinal summation mechanisms. These methods differentially activate classes of primary afferent fibers that mediate touch (large diameter A-beta fibers activated selectively by vibration, pressure, and electrical stimulation), pricking pain (thinly myelinated A-delta fibers activated by punctuate, heat, cold, and electrical stimulation), and diffuse, often burning pain (unmyelinated fibers activated by heat, blunt pressure, and electrical stimulation).

Measurement of Pain in the Clinic

Despite the complexity of pain, most clinical measures still treat it as a single dimension and use simple scales, such as a number from 1 to 10 or marking a VAS. These measures are used to assess pain in the clinic and often as the primary endpoint in extensive clinical trials of analgesic agents. Unlike other physical signs or symptoms, these measures are based on a person's evaluation and report of a private experience. Individuals may be instructed to evaluate the pain experience "right now" or asked to integrate the average pain over the last day, week, or other interval. Home pain diaries are used to capture fluctuations in pain and provide a more objective integration of pain over time. Concern for compliance has led to the use of electronic diaries that record the time of every entry and that provide a reminder function. On the other end of the objective–subjective continuum, measures of global perception of change in clinical trials are receiving attention because the judgments include subjectivity about overall quality of life.

The previous measures of pain magnitude are useful for the goals of testing analgesics and monitoring treatment efficacy. Other measures are useful for diagnosis and choice of treatment. One of the most popular is the McGill pain questionnaire (MPQ), which uses multiple verbal categories to assess pain intensity and unpleasantness, but also to evaluate the many varied qualities of pain sensation. These qualities include four types of pressure (compression, traction, lacerating, punctuate, described by words such as squeezing, pulling, cutting, and drilling), thermal qualities of heat, warm, cool, cold, and freezing, temporal qualities such as flickering or pounding, locations such as superficial or deep, and spatial qualities such as spreading or radiating. The profile for an individual is useful for diagnosis and for the selection of available treatments. This information can be supplemented by simple physical bedside QST measures or more extensive QST procedures if needed.

Clinical pain measurement is concerned with pain, the person, and pain impact. Psychosocial variables, such as depression and anxiety, accompany pain and need to be evaluated for treatment planning and treatment efficacy. Ability to follow a behavioral regime is an important factor, as are cognitive styles such as catastrophizing. This style refers to the tendency to view all possibilities in the darkest terms. Functional brain imaging studies have found brain activity related to depression in the insular cortex and amygdala, whereas catastrophizing is associated with multiple regions related to the anticipation of pain (contralateral medial frontal gyrus, ipsilateral cerebellum) and attention to pain (anterior cingulate cortex [ACC], bilateral dorsolateral prefrontal cortex), as well as to emotional (ipsilateral claustrum, interconnected to the amygdale) and motor responses (contralateral lentiform nuclei). These findings are particularly interesting because catastrophizing is a significant negative prognostic indicator of treatment success that fortunately can be treated by cognitive behavioral methods. In specific cases, this aspect of treatment may be as important as the treatment of pain pathology. One interesting question is whether the activity in the previously mentioned brain regions is altered by successful therapy, suggesting a mediating mechanism, or not altered by successful therapy, suggesting a risk factor for catastrophizing.

The measurement of pain depends on the goal of the measurement, including diagnosis, choice of treatment, efficacy of treatment, efficacy of a particular component of treatment (clinical trial), and basic research into underlying mechanisms. Another goal, not discussed, is the medicolegal determination of disability. The tools required depend on the goal, ranging from quick clinical procedures, extensive psychophysical methods, physiological measures, behavioral assessment, and interview and questionnaire assessment of pain and of the person.

Richard H. Gracely

See also Cutaneous Perception; Decision Making, Perceptual; Itch, Tickle, and Tingle; Pain: Cognitive and Contextual Influences; Pain: Placebo Effects; Pain: Treatments for Chronic; Perceptual Development: Touch and Pain; Phantom Limb; Theoretical Approaches

Further Readings

Chapman, C. R., Casey, K. L., Dubner, R., Foley, K. M., Gracely, R. H., & Reading, A. E. (1985). Pain measurement: An overview. *Pain, 22,* 1–31.

Gracely, R. H. (2005). Studies of experimental human pain. In S. McMahon & M. Koltzenburg (Eds.), *Wall and Melzack's textbook of pain* (5th ed., pp. 267–289). Philadelphia: Elsevier/Churchill Livingston.

Gracely, R. H., & Eliav, E. (2008). Measuring and assessing pain. In Y. Sharav & R. Benoliel (Eds.), *Orofacial pain and headache* (pp. 45–56). Edinburgh, UK: Mosby-Elsevier.

Melzack, R., & Katz, J. (2005). Pain assessment in adult patients. In S. McMahon & M. Koltzenburg (Eds.), *Wall and Melzack's textbook of pain* (5th ed., pp. 291–304). Philadelphia: Elsevier/Churchill Livingston.

Merskey, H., & Bogduk, N. (1994). *Classification of chronic pain* (2nd ed., prepared by the Task Force on Taxonomy of the International Association for the Study of Pain). Seattle, WA: IASP Press.

PAIN: COGNITIVE AND CONTEXTUAL INFLUENCES

Recent pain research has been greatly influenced by the gate control model of pain presented in 1967 by psychologist Ronald Melzack and neuroscientist Patrick Wall. They suggested that both physiological and clinical data, as well as everyday experience, run counter to the classical view that pain simply arises from overstimulation of the somatosensory system. Anatomical, physiological, and psychological evidence point to a complex interaction of both peripheral and central information in responding to noxious stimuli.

Melzack and Wall also noted that the amount of pain after an injury is greatly influenced by contextual factors. An athlete, soldier, or worker may suffer a severe wound yet not report pain until long after the event, likely because the individual's attention was focused upon some vital task when the injury occurred.

The gate control theory, emphasizing parallel ascending and descending effects within the nervous system, suggested that pain is not a single sensation. Rather, it has several distinct dimensions that Melzack and Kenneth Casey called (1) the sensory-discriminative system, (2) the motivational-affective system, and (3) the cognitive-evaluative system. Put simply, the first deals with identifying the location and intensity of the pain (how it feels), the second with our emotional response to that sensation (how it makes us feel), and the third with our interpretation and response to that event (how we think about it and act).

Consequently, pain research and management must deal with sensory, emotional, and cognitive mechanisms. This is reflected in pain measurement techniques, in functional imaging studies, in medical interventions, and in psychological approaches to pain reduction. This entry discusses some of the techniques used to measure pain, behavioral measures that illustrate how pain is influenced by cognition, and how pain is influenced by context effects.

Pain Measurement

Melzack and Warren Torgerson studied the basic dimensions of pain through an analysis of the English language. They came up with a list of 102 adjectives commonly used to describe elements of the pain experience. A group of subjects were asked to place these terms into categories, yielding three major classes. The classes were words that described the sensory qualities of the experience (such as their temporal, spatial, and thermal characteristics), the affective qualities (words such as *tiring, frightful,* or *wretched*), and evaluative ones that descried the overall character of the pain experience (such as *annoying, miserable,* or *unbearable*).

These terms were scaled for their intensity or severity, and an instrument called the McGill Pain Questionnaire (MPQ) was developed to measure the overall pain impression as well as its distinct components. The MPQ (now translated into more than 20 languages) has become widely used both for pain assessment (measuring, for example, changes in score after treatment) as well as diagnosis (because certain terms are used much more often by individuals suffering from one pain syndrome or another).

Words, of course, are only one means by which people can describe their level of pain or suffering. Numbers are another method; marks along a line (the visual analog scale) are a third. Richard Gracely devised several scales in which individuals were asked to report their pain intensity and distress separately. Although it is generally the case that increases in intensity are accompanied by increases in unpleasantness, that is not always so. Researchers sometimes remind people being asked to rate the two components that soft music that they dislike is likely to be markedly more unpleasant than loud music performed by a favorite band.

Groups of dental patients were administered diazepam, a tranquilizing antianxiety drug, and asked to report the intensity or unpleasantness of a noxious stimulus. The drug lowered only the affect or distress rating, leaving the intensity rating untouched. The dimensions did not change in unison. This is somewhat reminiscent of the descriptions given by psychiatric patients who had undergone prefrontal lobotomies. Those who had a long-standing pain problem reported that the "little pain" (the sensation) was still there but the "big pain" (the distress and suffering) had been attenuated or eliminated.

Functional Neuroimaging of Pain

Given the subjective nature of pain, it is impossible to obtain a truly "objective" measure of the experience. There is no "pain thermometer." In the final analysis, it is the individual's report of his or her experience that determines the evaluation of symptoms and response to them. Such reports are cortically determined. Although there are fascinating issues related to the receptors, nerve fibers, spinal pathways, and subcortical regions that convey signals generated by noxious inputs, it is the cortical regions that mediate, either within a complex regional neural matrix or through descending influences, such critical variables as context, expectations, pain beliefs, attention, and coping mechanisms in developing the pain percept.

Recent technical advances have permitted us to combine psychophysical and neuroimaging techniques in order to investigate the neural correlates of pain, making it possible to see how emotional and cognitive factors influence neural representations of the pain experience. Not surprisingly, a large number of brain regions have been implicated in the experience of even acute pain (chronic pain being even more complex). These include the primary and secondary somatosensory cortices as well as the insular, anterior cingulate, and prefrontal cortices (accompanied, sometimes, by motor cortices, the posterior parietal cortex, and the posterior cingulate), plus such subcortical regions as the thalamus, amygdala, hippocampus, cerebellum, and medulla.

Although the specifics change somewhat from study to study, at least partially due to methodological differences, there is general agreement that the sensory-discriminative aspects of pain perception, such as judgments about intensity, duration, quality, and location, are subserved within the SI and SII regions of the somatosensory cortex, and the motivational-affective and cognitive-evaluative aspects of pain tend to reflect activity in the insular cortex (IC), anterior cingulate cortex (ACC), and prefrontal cortex (PFC).

It is in these latter areas that unpleasantness, suffering, and evaluation of the pain experience are mediated. The IC is seen as important in emotional responses to noxious input, and the ACC is considered to play a critical role in both emotion and cognition, particularly with respect to anticipation and expectation of pain and the interplay of attention and response planning. It is here where activity is reduced by hypnotic suggestions for the reduction of pain.

The PFC may act as a sort of supervisory attention system or "pain control center," integrating emotion and such cognitive components as perceived control over a painful event. Interestingly, not all studies of brain activation by noxious signals show activity in the PFC in healthy subjects, although chronic pain conditions are often associated with increases there, as well as in the somatosensory cortex and the ACC.

Activation of these regions does not necessarily require noxious stimulation. A series of recent studies have shown that observing films showing another individual receiving painful stimuli may trigger a sort of empathic reaction, causing activity in the ACC and IC not unlike that created by painful stimulation itself. In another study, when subjects observed pain from the faces of chronic pain patients, activations in the IC, ACC, and parietal lobe in the observer's brain correlated with their estimates of the intensity of observed pain. Even imagining pain through hypnotic induction activates the thalamus, ACC, IC, PFC, and the parietal cortex.

Other cognitive tasks can modulate the cortical response to noxious inputs. The anticipation of pain, perhaps by creating hypervigilance to impending threat, increases activity in the emotion-processing network consisting of the IC, ACC, and amygdala. Just the task of providing a discomfort rating increases brain activity in the cortical pain matrix. Such effects are bidirectional. Many of these regions show decreases in activity when subjects receive placebo analgesia. Particularly noteworthy is the finding that these changes in neural response occur during the stimulus itself, rather than somewhat later when pain reports are made.

Behavioral Studies of Pain and Cognition

Cognitive effects can modulate pain in two directions. That is, such factors as anticipation of pain, attention, anxiety, and maladaptive thoughts increase pain and suffering, yet distraction, hypnotic suggestion, placebo administration, cognitive therapies, and induction of positive mood states can reduce both experimentally induced pain and reports of pain and distress created by chronic pain syndromes.

In laboratory studies, cognitive strategies such as engaging in mental imagery, interpreting the stimulation as something else, or diverting attention to environmental cues have often been shown to enhance pain threshold or tolerance or reduce pain ratings to a constant stimulus. The literature in this field is murky, likely due to large differences in pain induction methods (generally various forms of intense heat, cold, pressure, or electrical pulses), instructions, content, and nature of the coping strategies used and measures taken. There are differences even in the general nature of attentional control; distraction can diminish pain, but so too can deliberate focus upon the part of the body receiving stimulation, in which subjects are told to attend to the information coming from the site while ignoring the emotional aspects (redefinition).

It is not only the case that attention affects pain; the reverse is also true. Pain often serves as a warning of impending danger, and neglect of this signal is highly maladaptive. In the laboratory, pain has been seen to significantly disrupt performance on various attentionally demanding tasks involving perception, learning, memory, vigilance, and movement. Likewise, pain patients often show impairment in everyday cognitive activities (things such as recall of people, places, and activities or the use of spoken and written language), a phenomenon some fibromyalgia patients label "fibrofog."

Cognitive response styles characterize how individuals deal with potentially threatening situations. Some people are prone to use adaptive thought patterns (coping). Others use maladaptive coping styles (catastrophizing) in which they expect negative outcomes, show high levels of fear or hypervigilance to threat, misinterpret events, and are impaired in their ability to divert attention away from their pain. Psychological tools (the Coping Strategies Questionnaire or the Pain Catastrophizing Scale) can measure the tendency to engage in adaptive or maladaptive thought patterns; others (the Pennebaker Inventory of Limbic Languidness and the Modified Somatic Perception Questionnaire) assess attention to physical sensations and symptoms.

Contextual Effects in Pain Judgments

Pain judgments are not made in a vacuum. There is considerable evidence that descriptions and ratings of pain are influenced by contextual information about the environment, circumstances, and anticipation of the noxious stimulus. Dental and medical students are advised to notify patients just before administration of a procedure (such as an injection) is likely to become painful, so that the patient is able to prepare for the event. Even knowledge that one can influence the course of the clinician's action, through finger signs or grunts, helps patients to attenuate both negative affect and pain.

Gary Rollman found that pain ratings are influenced by other stimuli in the presentation set. A stimulus will be judged as less painful when presented in the same session as a stronger one than it is when paired with a weaker stimulus. This adaptation level effect, which is similar to ones found in other modalities, highlights the fact that pain judgments are relative rather than absolute.

In the case of pain patients, judgments regarding the painfulness of experimentally induced discomfort are often lower than those provided by individuals who are pain free. Endogenous pain serves as an anchor or comparison point by which newly added stimuli are judged. This adaptation level model is also applicable to at least some instances of a phenomenon known as diffuse noxious inhibitory controls (DNIC), in which strong, tonic pain at one part of the body reduces the response to a phasic pain stimulus presented elsewhere. Seen in both humans and lower animals, the DNIC effect is at least partly mediated by endogenous opiates, but cognitive comparisons are also fundamental factors.

Mental states induced by placebo instructions are also capable of releasing the body's endogenous opiates, such as endorphin. Thoughts and expectations have physiological effects; placebos are not biologically inert. For example, dental patients who are given placebos show reductions in pain ratings to induced stimuli, which are reversed by the administration of naloxone, an opiate antagonist drug that blocks the receptor sites where endorphins bind. Moreover, neuroimaging studies have revealed that decreases in pain ratings after administration of a placebo are accompanied by decreases in activity in the ACC, IC, and PFC.

Reflections

The literature makes clear that pain is not simply an overstimulation of the central nervous system. The richness of the pain experience, the multiplicity of neural sites that respond to noxious signals, and the amplification or diminution of pain created by emotions, thoughts, and expectations make clear that pain is essentially constructed by a cascade of peripheral and central events involving ascending and descending neural networks. This distributed pain network evaluates incoming information within both a bottom-up and a top-down context determined by memories of previous events, current emotional states, pathology, genetics, and many cognitive variables. There has been considerable success in utilizing cognitive-behavior therapies, which address thoughts, beliefs, appraisals, attitudes, and coping strategies, to ameliorate both acute and chronic pain states. While cognitions can enhance pain and suffering, they can also be utilized to diminish those distressing conditions.

Gary B. Rollman

See also Attention: Cognitive Influences; Brain Imaging; Emotional Influences on Perception; Multimodal Interactions: Pain–Touch; Pain: Assessment and Measurement; Pain: Neuromatrix Theory; Pain: Physiological Mechanisms; Pain: Placebo Effects

Further Readings

Apkarian, A. V., Bushnell, M. C., Treede, R. D., & Zubieta, J. K. (2005). Human brain mechanisms of pain perception and regulation in health and disease. *European Journal of Pain, 9,* 463–484.

Eccleston, C., & Crombez, G. (1999). Pain demands attention: A cognitive-affective model of the interruptive function of pain. *Psychological Bulletin, 125,* 356–366.

Lautenbacher, S., Prager, M., & Rollman, G. B. (2007). Pain additivity, diffuse noxious inhibitory controls, and attention: A functional measurement analysis. *Somatosensory & Motor Research, 24,* 189–201.

Rollman, G. B. (1992). Cognitive effects in pain and pain judgments. In D. Algom (Ed.), *Psychophysical approaches to cognition. Advances in psychology* (Vol. 92, pp. 517–574). Amsterdam: North-Holland.

Tracey, I., & Mantyh, P. W. (2007). The cerebral signature for pain perception and its modulation. *Neuron, 55,* 377–391.

Wiech, K., Ploner, M., & Tracey, I. (2008). Neurocognitive aspects of pain perception. *Trends in Cognitive Sciences, 12,* 306–313.

PAIN: NEUROMATRIX THEORY

The *neuromatrix theory* proposes that the anatomical substrate of the body-self is a widespread network of neurons that consists of recurrent loops between the thalamus and the cerebral cortex. The traditional theory of pain, which evolved during the early 20th century, holds that pain sensation is produced by a direct-line spinal cord pathway from "pain receptors" in the body to a "pain center" in the cerebral cortex. Research based on this theory focuses on acute pain evoked by noxious stimulation and has revealed complex physiological mechanisms at every level of the pathway from receptors to the cortex. However, it has failed to explain types of chronic pain, such as relentless backaches in the absence of any pathology or phantom limb pain that persists for decades after the amputation of a limb.

Phantom limb pain after amputation of a limb reveals the powerful role of the brain in chronic pain. A high-level cordectomy—total removal of several segments of spinal cord so that sensory information from the pelvis and legs is unable to arrive at the brain—does not stop intense pain in the phantom half of the body. The extraordinary reality of painful phantom limbs indicates that the brain does more than detect and analyze sensory inputs; it generates perceptual experience even in the absence of external inputs. We do not need a body to feel pain or a physical injury to elicit pain. The brain can generate both experiences. These conclusions are the basis of a new theory of pain: the neuromatrix theory, which is described in this entry.

The spatial distribution and synaptic links of the body-self neuromatrix are initially determined genetically and are later sculpted by sensory inputs. The loops diverge to permit parallel processing in different components of the neuromatrix and converge in the thalamus to permit interactions among the output products of processing. The repeated *cyclical analysis and synthesis* of nerve impulses through the neuromatrix imparts a characteristic pattern: the *neurosignature.* The neurosignature of the neuromatrix is imparted on all nerve impulse patterns that flow through it as a result of the patterns of synaptic connections in the entire neuromatrix. All inputs from the body undergo cyclical analysis and synthesis so that characteristic firing patterns are impressed on them in the neuromatrix. Portions of the neuromatrix are specialized to process information related to major sensory events—such as injury, temperature change, and stimulation of erogenous tissue—and are labeled as neuromodules, which impress subsignatures on the larger neurosignature.

At any instant in time, millions of nerve impulses arrive at the brain from all the body's sensory systems, including the proprioceptive and vestibular systems. How can all this be integrated to a unity of experience? As a result of the constant, spontaneous neural activity in the brain, the body-self neuromatrix produces a continuous message that represents the whole body in which details are differentiated within the whole as inputs come into it. The neuromatrix, then, is a template of the whole, which provides the characteristic neurosignature pattern for the whole body as well as subsets of signature patterns from neuromodules that relate to events at (or in) different parts of the body.

The experience of the body-self involves multiple dimensions—sensory, affective (emotional), and cognitive (see Figure 1). The sensory dimensions

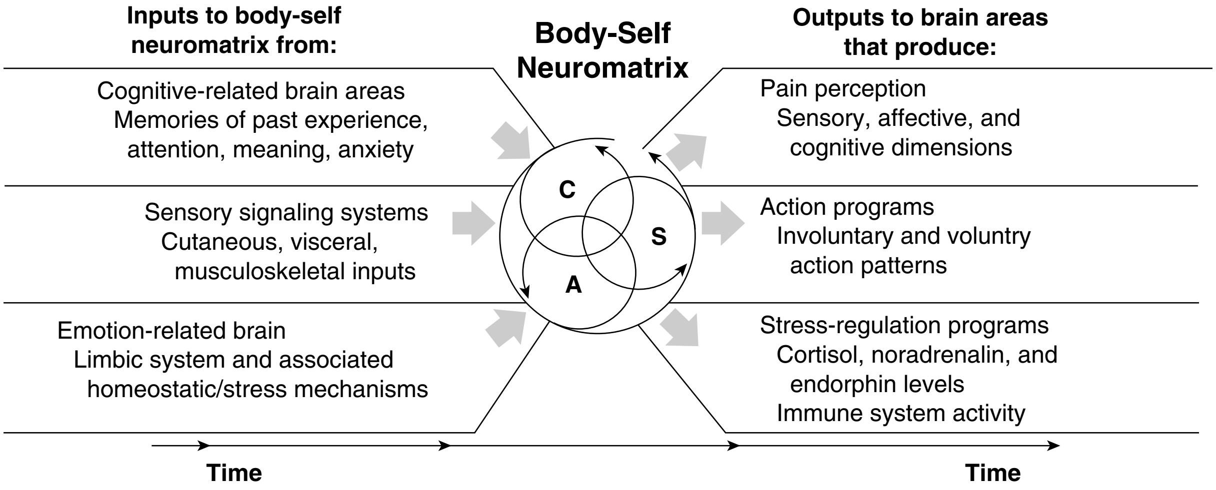


Figure 1 Factors That Contribute to the Patterns of Activity Generated by the Body-Self Neuromatrix

Source: Melzck, R. and Wall, P.D. The challenge of pain. Penguin Books, London. Reprinted with a new introduction, 2008. The figure appears on p. xv.

Notes: The neuromatrix is composed of sensory (S), affective (A), and cognitive (C) neuromodules. The output patterns from the neuromatrix produce the multiple dimensions of pain experience, as well as concurrent behavioral and homeostatic responses.

are subserved, in part at least, by portions of the neuromatrix that lie in the sensory projection areas of the brain; the affective dimensions are subserved by areas in the hypothalamus and limbic system. The major cognitive components of experience, such as attention and expectations, are subserved by the frontal and parietal cortical areas.

After inputs from the body undergo partial transformation in the body-self neuromatrix, the neuromatrix for action patterns is activated concurrently with the neuromatrix for experience. During the time course of cyclical analysis and synthesis, several possible behavior patterns may be activated and serially eliminated until one particular pattern emerges as the most appropriate for the situation at the moment. In this way, inputs and outputs are synthesized in parallel, thereby permitting a smooth, continuous stream of perceptual experience and action patterns.

The continuous, spontaneous activity in the body-self neuromatrix provides an explanation of our constant awareness of our body and our sense of self. That a neuromodule exists to produce awareness of our "self" in space and time is evident from the fact that lesions in specific areas of the brain—from the midbrain reticular core to the parietal cortex—produce a loss of the sense of self for half of the body as well as the space and its contents surrounding that side of the body. We take the "sense of self" for granted, but it is generated by the body-self neuromatrix and can be abolished by specific brain lesions.

Figure 1 presents the concept of the "neuromatrix" that underlies the complexity of pain. Neural programs that generate the awareness of pain evolved in the body-self neuromatrix as a powerful mechanism to produce escape from predators that inflict injury. At the same time, other programs evolved to activate the sympathetic nervous system to produce a return to homeostasis after injury or infection. However, these programs also activate stress-regulation mechanisms that may cause stress-related disease and pain. It is possible that neural programs that evolved to maintain homeostasis may, instead, go awry and produce chronic disease and pain.

Ronald Melzack and Joel Katz

See also Cutaneous Perception; Pain: Cognitive and Contextual Influences; Pain: Physiological Mechanisms; Pain: Treatments for Chronic; Phantom Limb

Further Readings

Dostrovsky, J. O., & Craig, A. D. (2006). Ascending projection systems. In S. B. McMahon &

M. Koltzenburg (Eds.), *Wall and Melzack's textbook of pain* (5th ed., pp. 187–204). Philadelphia: Elsevier/Churchill Livingston.

Melzack, R. (1989). Phantom limbs, the self and the brain (The D.O. Hebb Memorial Lecture). *Canadian Psychology, 30,* 1–14.

Melzack, R. (2005). Evolution of the neuromatrix theory of pain. The Prithvi Raj Lecture. *Pain Practice, 5,* 85–94.

PAIN: PHYSIOLOGICAL MECHANISMS

In the late 1960s, John F. Hahn, a sensory generalist in the tradition of Henri Piéron and Frank Geldard, taught that pain is not only a sensation in search of a stimulus, but in search of a receptor as well. In the years since those discussions, understanding of the mapping of stimulus on receptor to sensation has not changed much. There is no question that pain is a dramatic and attention-grabbing event when it occurs—when skin temperature is too high or too low, when we taste or touch chemicals, such as capsaicin (from chilies) or when we stub our toe—and then again a few seconds later, but perceived differently—duller and less sharp. But one could argue that the most interesting aspect of pain is when it doesn't occur and yet seems like it should. Phenomena such as phantom limb pain, the wounded soldier who continues to fight unaware of an injury, or the "anesthetic" effect of clenched fists in response to stubbing a toe underscore the fact that pain is a separate, complex sensory-perceptual experience. These also illustrate that the experience of pain is strongly subject to nonsensory ("top-down") central influences, like affective state and peripheral modulating influences. Perhaps more than any other sensory modality, it can't easily be explained by stimuli activating a pain receptor and being transmitted to the brain, challenging straightforward physiological analyses. But physiological mechanisms must underlie at least the initial triggering of the experience. This entry begins by placing the search for physiological mechanisms in historical context and then describes the results of contemporary research on physiological mechanisms.

Pain Is Different From Other Tactile Sensations

The perception of pain is different from other sensory experiences, and the physiology of "pain receptors" has been similarly difficult to define. At least tactually, pain is not just "very intense touch." Historically, there had been a debate whether painful stimuli were just examples of very intense energy of some type (heat, pressure, etc.), reported by our touch receptors responding to these extremes on a continuum from normal and innocuous to painful, injurious, and noxious. The alternative view has been that there were specific receptors, called *nociceptors,* for the stimulus extremes. Observations described in the following text relating certain fiber types to stimuli described as "painful" support the latter notion. The brain, which reports the presence of pain or inhibits our appreciation of this illusive experience, can be explored surgically without resulting in pain, as Wilder Penfield was able to demonstrate in his explorations of cortical function during procedures when the patient was awake, because there are no pain receptors in the brain itself. Possessing only free nerve endings, somehow the cornea of the eye is able to distinctly report touch, temperature, and pain. (Bruce MacIver and Darrell Tanelian do point out that there is a distinction in the way in which the nerves branch out in the cornea, but there is no other apparent specialization in the endings.) So pain involves both affective (central) and sensory (peripheral) components. Other entries discuss some of the cognitive determinants of pain, its measurement, and current theories involving central nervous system pathways and interactions among them. This entry discusses the physiological bases of pain, particularly in the responses of specific peripheral receptor systems, called nociceptors.

Fiber Types Provide Evidence for Separate Pain Pathways

Historically, there have been a number of approaches to understanding the physiological basis of this perceptual experience, some of which have led to fairly aggressive experiments. For example, Stanley Finger relates how Sir Henry Head, in the 1890s, frustrated at the impatience of his students, cut the radial and other sensory nerves conducting signals from the skin toward the brain in his own arm to observe the course of recovery of the cutaneous senses. Head noted that initially, regardless of the stimulus—cold, warm, or touch—pain would also be evoked. He argued that this phenomenon indicated that pain must be subserved by the finest of sensory nerve endings that regrow most rapidly, rather than the larger fibers. We experience something similar after our arm or leg "falls asleep" (paresthesia): The "pins and needles" during recovery reflect disorganized return of our sensations. This distinction, between the response of fine fibers and that of the larger fibers, would continue through the literature. This division fell into line with observations by Magnus Blix (ten years earlier), who explored the sensitivity of punctuate ("pointlike") spots on the skin with electrical and thermal stimuli and reported that those regions sensitive to pressure, warmth, cold, and pain did not appear to overlap.

It was only a few years later that Herbert Gasser and Joseph Erlanger conducted the earliest nerve block studies in which they induced ischemia in the arm with a tourniquet, starving the nerves by cutting off their blood/oxygen supply. This created an anesthetic state in the limb. Although the senses of touch were perceived to disappear quickly, sensitivity to pain remained—at least for a little while longer, again supporting the notion that these submodalities were served by separate underlying systems. They were further able to demonstrate that electrical signals recorded from the whole nerve showed distinctly different speeds of conduction (indicating activity in nerve fibers of distinctly different sizes). These were related to the serial disappearance of specific sensations, further linking the sense of touch to the faster (larger) fibers and pain to the slower (smaller) fibers. In other studies, local anesthetics like lidocaine or Novocain have been found to block the electrical activity in the smallest fibers first and the larger ones later. Perceptually, use of these anesthetics leads to a progressive loss of pain sensitivity *before* cold and touch disappear—just the opposite order to that found with ischemia. Similar dissociations among skin sensations are seen as symptoms in affected spots with syphilis, where pain and itch are lost, and with leprosy (Hansen's disease), where only touch and temperature are lost.

Finally, medical reports regarding spinal cord injury have shown that this functional separation is followed into the spinal cord—with one set of tracts carrying non-noxious tactile information, and a separate one carrying pain information (the lateral spinothalamic tract). Within the spinal cord, possibilities of interactions exist—not only within the fiber group containing pain information, but with descending controls from cortical structures as well. Therapists have been able to take advantage of these functional spinal separations between touch and pain and the existence of modulating interactions within the spinal cord in attempts to treat intractable pain. For example, surgeons can actually sever the pain pathway (in a procedure called anterolateral cordotomy, first performed in 1911 by Edward Martin and Artur Schüller), leaving touch intact. In other cases, stimulating modulating areas in the spinal cord electrically (transcutaneous electrical nerve stimulation) can reduce pain. Neither is completely successful, for reasons including anatomical variability as well as perceptual and cognitive factors.

Gate-Control Theory

Additional evidence supports a further separation of function by fiber type within the group of small fiber pain systems. The first is made up of the smallest of the A fibers (called A delta or Aδ). Larger A fibers appear to be responsible for "normal" touch. The whole population of A fibers is characterized as being myelinated, which means that the nerve fibers are "wrapped" by a particular type of cell, called Schwann cells, that have the effect of speeding nerve conduction. The second type consists of unmyelinated C fibers, and because of their small size and lack of myelination they conduct nerve impulses at a fraction of a meter per second—several times slower than the Aδ fibers. The differences between these two fiber types, especially their conduction speeds, underlies the experience of "double pain" felt, for example, when we stub a toe. Edward Perl and Lawrence Krueger give a brief history of the percept of a fast sharp pain followed later by a deep dull pain. Perl and another colleague, Bruce Lynn, further review studies of these two groups of small fibers tested with microneurographic techniques. In these studies, humans were asked to report sensations produced when the fibers were electrically stimulated, providing the important link between the neural events that occur deep in the skin and perceptual experience. When a fiber in the forearm was electrically stimulated through the microelectrode, the subject would feel "something" at the distal termination of that nerve, for example, in the fingertip. This percept is described as a "referred" sensation. Two different kinds of pain sensation were described at the referred sites: "sharp, pricking pain" was felt when Aδ fibers were stimulated, whereas the sensations resulting from C fiber stimulation depended on the type of skin. If the referred site was in glabrous (smooth) skin, like the palm of the hand, stimulation was felt as dull, whereas for referred sites in hairy skin, it was felt as burning.

The differentiation between these two types of "pain" fiber types was used as the basis of one of the more useful physiological models to explain the way in which the perception of pain produced by the incoming ("afferent") neural activity was modulated by descending ("efferent") information from other sites, including the central nervous system. This efferent control of the incoming signals is proposed to occur in the spinal cord, and is described in the Melzack-Wall gate-control theory. While responses to noxious stimuli enter the "gate" through the small fibers, efferent influences can close it, modulating or reducing the pain. This model has more recently been elaborated into the neuromatrix theory. As James Craig and Gary Rollman relate, this theory attempts to bring together the information from the peripheral nervous system with the modulating influences from higher cortical levels to determine whether our response to aversive or noxious stimuli is perceived as pain, but this linkage is far from straightforward or well understood. For example, Mark Hollins and his colleagues studied how vibration on the forearm modulates the noxiousness of a brief burst of laser-produced radiant heat. In order to separate the changes in sensitivity from the possibility of changes in bias, he had to use an extended application of signal-detection methodology to show that over a wide range of frequencies and intensities, vibration could reduce pain mediated by Aδ fibers. This laboratory demonstration quantifies the everyday experience that occurs when one is able to "mediate" the pain from a minor injury at one body site (e.g., being kicked in the shin) by squeezing one's hands or rubbing the overlying skin. So it should not be surprising that if some types of pain can be controlled in this simple case, there can be more profound descending central influences—including the individual's attentional and affective states—on other types of pain.

The Search for Pain Receptors

The question remains, though, whether these nerve fibers, associated with painful sensations, are linked to specific structures in the periphery that exclusively encode painful (as opposed to nonnoxious) stimuli. As elusive as the search for receptors for the other submodalities of the skin's senses (e.g., vibration or temperature) has been, the search for "pain" receptors has been even more difficult. What has been established is that there are, in fact, neural systems that are truly nociceptive. But what underlies those peripheral sensory spots, some of which were described by Blix as only generating a painful sensation when pricked or pinched or burned? Joe Stevens and Barry Green suggested that these sites probably represented the subcutaneous presence of individual low-threshold receptors tuned to those particular types of energy, or dense aggregates of similarly sensitive receptors, or even regions where more complex neural coding might be taking place. There does appear to be an underlying physiological difference in the receptor sites for different intensities of tactile stimuli: When mechanoreceptors or thermoreceptors are stimulated with levels of energy that humans describe as painful, their neural response either stops altogether or the response is just a bit higher than their response to lower energy levels. For example, although "warm" thermoreceptors will respond to skin temperatures approaching the painful range of about 41 to 43° Celsius, their firing rate is simply higher than that evoked by the normal-to-warm temperatures of 35 to 40°. However, heat-sensitive nociceptors will only respond to the hot-to-painful potentially injurious range of temperatures from about 43 to 50° (or higher). In their comprehensive description, Bruce Lynn and Edward Perl detail the receptor systems that seem to subserve the basic submodalities of information that one would classify as painful. But in their description, only one type of noxious stimulus has been tied to a particular gross neural structure—the morphology (form and structure) of the A mechano-nociceptor nerve ending appears to have a unique appearance in the skin. None of the other nociceptor types, A or C fiber, mechanical, thermal, or chemical, appear to have particular terminal structures or organization that would best suit them to process potentially injurious stimuli. However, at the level of the nerve membrane, recent research has explored specialization of specific receptors that open or close ion channels in response to specific noxious stimuli. The monograph edited by Uhtaek Oh describes a number of these membrane-level mechanisms in nociceptive nerve fibers and even discusses the cloning of receptors specific to particular noxious stimuli, such as capsaicin.

Characteristics of Pain Receptors

Researchers now know that there are unique nociceptors in a wide range of structures and organs in the body—the teeth, the cornea of the eye, the heart, the testes, the uterus, and others. However, this section deals primarily with those described in the skin, our largest organ. Lynn and Perl can be referred to for details on some of these other sites. Those in the skin have been best studied, and their functions and sensitivity to noxious stimuli seem to be divided between the small Aδand C fibers. Aδ nociceptors are of two types: those sensitive to mechanical stimulation (like pinching or hard probing) and those sensitive to heat or chemicals. The mechano-sensitive nociceptors tend to have multiple receptive fields, much like cutaneous type II SA receptors in hairy skin. Those responsive to heating (in the 40–50° range) seem to be more common, with some firing vigorously in the very high (>50°) range, particularly with long-duration stimuli. It is likely that some human responses to very high heat may be mediated by these because response times are too rapid to be mediated by the other population of heat-sensitive nociceptors, the C fiber temperature-sensitive ones.

C fiber nociceptors conduct more slowly and appear to be the source of that duller, slower pain in "double pain" experience. They typically have single well-demarcated receptive fields, in contrast to the Aδ pattern. Furthermore, they are often polymodal. That is, they respond in a regular fashion (increasing their neural firing rate as the stimulus intensity increases) to very strong mechanical and high-temperature noxious stimuli as well as to chemical irritants like capsaicin, acids, histamine, or mustard oil. Some have also been reported to respond to extreme cold (about 15°), in contrast to cool receptors that respond to nonaversive cold stimuli produced by as little as a 1° reduction from normal skin temperature (about 35°). C fiber nociceptor populations have also been found that are only sensitive to noxious mechanical (but not thermal) stimuli, only to high temperatures but not mechanical pressure, to chemical stimulation only, and, rarely, only to cold stimuli.

A number of behaviors of nociceptors have been observed at the physiological level (nerve recordings) that are mirrored in behavioral experience. One of the more interesting is C fiber sensitization. Some C fiber nociceptors don't respond until a prior "sensitizing" event occurs. In particular, joint nociceptors have been found that are only active if there is local inflammation. Similarly, stimulation with chemical irritants can induce sensitivity in other C fiber nociceptors to a previously nonadequate stimulus: They become sensitive to classes of noxious stimuli to which they would not previously respond. So these findings suggest that prior injury can activate the system and increase its responsiveness. Conversely, C fiber suppression occurs in other cases where C fiber nociceptors show dramatically reduced responses a short time after aversive (e.g., heat) stimuli. Not only does this suppression last as long as five minutes after the inducing stimulus, but it is directly related to the intensity of the inducer. From an everyday perceptual standpoint, it is interesting to note that the pain resulting from the chemosensory irritation that occurs when eating chili pepper recipes can be attenuated by pausing for five minutes until the pain subsides. Barry Green has found that continuing to eat afterward will be less aversive, but depends on the intermission; the suppression effect will not occur if capsaicin is constantly present in the mouth. These effects, both suppression as well as sensitization, appear to be peripheral in origin, occurring at the end of the fibers that are responsible for the initial encoding of nociceptive stimuli that may be perceived as pain by the central nervous system, and have remarkable perceptual parallels.

Roger W. Cholewiak
and Steven A. Cholewiak

See also Cutaneous Perception: Physiology; Pain: Cognitive and Contextual Influences; Pain: Neuromatrix Theory; Pain: Placebo Effects

Further Readings

Brown, J. (1973). *Best & Taylor's physiological basis of medical practice* (9th ed.). Baltimore: Williams and Wilkins.

Craig, J., & Rollman, G. (1999). Somesthesis. *Annual Review of Psychology, 50,* 305–331.

Finger, S. (1994). *Origins of neuroscience: A history of explorations into brain function.* New York: Oxford University Press.

Hollins, M., Roy, E. A., & Crane, S. A. (2003). Vibratory antinociception: Effects of vibration amplitude and frequency. *The Journal of Pain, 4*(7), 381–391.

Iggo, A. (1982). Cutaneous sensory mechanisms. In H. B. Barlow & J. D. Mollon (Eds.), *The senses* (pp. 369–408). Cambridge, UK: Cambridge University Press.

Lynn, B., & Perl, E. (1996). Afferent mechanisms of pain. In L. Krueger (Ed.), *Pain and touch* (pp. 179–211). New York: Academic Press.

MacIver, B., & Tanelian, D. (1993). Structural and functional specialization of Aδ and C fiber free nerve endings innervating rabbit cornea1 epithelium. *The Journal of Neuroscience, 13*(10), 4511–4524.

Oh, U. (2006). *The nociceptive membrane.* New York: Academic Press.

Penfield, W., & Rasmussen, T. (1950). *The cerebral cortex of man.* New York: Macmillan.

Perl, E., & Kruger, L. (1996). Nociception and pain. In L. Krueger (Ed.), *Pain and touch* (2nd ed., pp. 179–211). New York: Academic Press.

Roy, E., Hollins, M., & Maixner, W. (2003). Reduction of TMD pain by high-frequency vibration: A spatial and temporal analysis. *Pain, 101,* 267–274.

Stevens, J., & Green, B. (1996). History of research on touch. In L. Krueger (Ed.), *Pain and touch* (pp. 1–23). New York: Academic Press.

PAIN: PLACEBO EFFECTS

The *placebo effect* is the reduction or the disappearance of a symptom when an inert treatment (the placebo) is administered to a subject who is told, and indeed believes and expects, that it is an effective therapy. Most of the knowledge about its mechanisms comes from the field of pain, thus placebo analgesia is currently the most investigated model. However, other systems and apparatuses, such as the motor, immune, and endocrine systems, are emerging as interesting models. Although the placebo effect has so far been considered a nuisance in clinical research when a new treatment has to be tested, it has now become a target of scientific investigation to better understand the physiological and neurobiological mechanisms that link a complex mental activity to different functions of the body. Usually, in clinical research the term *placebo effect* refers to any improvement in the condition of a group of subjects that has received a placebo treatment. Conversely, the term *placebo response* refers to the change in an individual caused by a placebo manipulation. However, today these two terms are used interchangeably.

The placebo effect is basically a context effect, whereby the psychosocial context around the patient plays a key role. For example, the therapist's words, the sight of complex machines, and other sensory inputs that tell the patient that a treatment is being performed, all represent important factors in the occurrence of a placebo response. In the case of pain, this psychosocial context is capable of modulating pain perception. This is the reason why the placebo effect is currently a useful model for understanding the complex psychological modulation of pain. This entry describes the identification and mechanisms of the placebo effect, as well as the nocebo effect (a placebo effect in the opposite direction).

Identification of the Placebo Effect

The investigation of the placebo effect is full of pitfalls and drawbacks because, in order to identify a real psychobiological placebo response, several other phenomena have to be ruled out. For example, most painful conditions show a spontaneous temporal variation that is known as natural history. If subjects take a placebo just before their discomfort starts decreasing, they may believe that the placebo is effective, although that decrease would have occurred anyway. Clearly, this is merely a misinterpretation of the cause-effect relationship. Another example is regression to the mean, a statistical phenomenon whereby individuals, after reporting severe pain at an initial clinical assessment, tend to report lower levels of pain at a second assessment. A further source of confusion is represented by signal detection theory, whereby errors in the detection of ambiguous signals may occur. For example, the subject may believe that a pain reduction has occurred, a false positive, although no real reduction is present. It also happens that a co-intervention actually is responsible for the reduction of a symptom, such as the analgesic effect following the mechanical insertion of a needle to inject an inert solution. All these examples show that, although an improvement may occur after the administration of a placebo, the placebo is not necessarily the cause of the effect that is observed.

Because all these phenomena are sometimes difficult to identify, the mechanisms of the placebo response must be investigated under strictly controlled experimental conditions. For example, in order to rule out spontaneous remission, a group taking the placebo is compared to a group receiving no treatment, the latter giving information on the natural course of the symptom. The difference between the placebo group and the no-treatment group represents the real psychobiological placebo response.

Mechanisms

The placebo effect involves both cognitive factors and conditioning mechanisms. For example, the deceptive administration of a placebo treatment can lead the subjects to believe that the treatment is effective, so as to induce positive expectations about the therapeutic outcome. Indeed, several studies show that different verbal instructions lead to different expectations and thus to placebo responses of different magnitude. The context around a therapy may therefore act through more than only expectation and conscious anticipatory processes. In fact, the placebo response is sometimes a conditioned response due to repeated associations between a conditioned stimulus (e.g., shape and color of pills) and an unconditioned stimulus (the active substance inside the pill). In this case, it is the context itself that is the conditioned stimulus. However, even by considering a typical conditioning procedure, it has been shown that a conditioned placebo analgesic response can result from conditioning but is actually mediated by expectation. In other words, conditioning would lead to the expectation that a given event will follow another event, and this occurs on the basis of the information that the conditioned stimulus provides about the unconditioned stimulus.

There is not a single placebo effect but many effects, with different mechanisms at work in different circumstances. Sometimes cognitive mechanisms are more important than conditioning, and vice versa. For example, verbally induced expectations of either analgesia or hyperalgesia completely eliminated the effects of a conditioning procedure in experimentally induced pain. By contrast, verbally induced expectations of either increase or decrease of hormones do not have any effect on their secretion. However, if a pharmacological preconditioning is performed by means of repeated administrations of a hormone-stimulating drug, conditioned hormonal placebo responses can be induced. These findings suggest that placebo responses are mediated by conditioning when unconscious physiological functions, like hormonal secretion, are involved, whereas they are mediated by expectation when conscious physiological processes, like pain, come into play. Thus, the placebo effect seems to be a phenomenon that can be learned either consciously or unconsciously, depending on the system that is involved (e.g., pain or hormone secretion).

There is now compelling experimental evidence that in some circumstances placebo-induced analgesia is mediated by endogenous opioid systems—opioid systems that originate naturally in the body. In fact, several studies found that placebo analgesia is reduced by the opioid antagonist, naloxone. In addition, some in vivo receptor binding studies, whereby the activity of neurotransmitter receptors is assessed, found the activation of opioid neurotransmission in some brain regions, like the anterior cingulate cortex, the dorsolateral prefrontal cortex, the insula, and the nucleus accumbens. Moreover, the cholecystokinin (CCK) antagonist, proglumide, has been found to enhance the placebo analgesic effect, which indicates that CCK has an inhibitory role in placebo-induced analgesia. The placebo analgesic response is thus the result of the balance between endogenous opioids and endogenous CCK.

Placebo analgesia is not always mediated by endogenous opioids. In fact, if the placebo response is induced by means of either strong expectation cues or morphine preconditioning, it can be blocked by the opioid antagonist naloxone. Conversely, if the placebo response is induced by means of prior conditioning with a non-opioid drug, like ketorolac, it is naloxone insensitive. Today we know that specific placebo analgesic responses can be obtained in different parts of the body, and that these responses are naloxone reversible, which suggests that the placebo-activated endogenous opioid systems have a somatotopic organization. In other words, opioid-mediated placebo responses occur only in those parts of the body where a placebo cream had been applied.

The investigation of placebo analgesia by means of brain imaging techniques found that similar regions of the brain, such as the anterior cingulate cortex, are affected by both a placebo and a narcotic drug, which indicates a related mechanism in placebo-induced and opioid-induced analgesia. In addition, during the anticipatory phase of the placebo analgesic response (i.e., when the subject expects the analgesic effect), an activation of the dorsolateral prefrontal cortex, orbitofrontal cortex, superior parietal cortex, periaqueductal gray, and other frontal regions occurs, suggesting the activation of a cognitive-evaluative network just before the placebo response.

Brain imaging studies have also shown the involvement of dopamine in the nucleus accumbens in placebo analgesia. As the nucleus accumbens is involved in the neuronal circuitry of reward mechanisms, these studies suggest that an important mechanism in placebo responsiveness might be reward. In the case of the placebo effect, the reward is represented by the clinical benefit.

The Nocebo Effect

The nocebo effect, or response, is a placebo effect in the opposite direction, whereby expectation of pain increase may induce a hyperalgesic effect. In this case, anticipatory anxiety plays a fundamental role, as the induction of a nocebo response represents a stressful and anxiogenic procedure. The term *nocebo* (“I shall harm”) was introduced in contraposition to the term *placebo* (“I shall please”) by a number of authors in order to distinguish the pleasing from the noxious effects of placebo. It is important to stress that the study of the nocebo effect relates to the negative psychosocial context surrounding the treatment, and its neurobiological investigation is the analysis of the effects of this negative context on the patient’s brain and body. As with the placebo effect, the nocebo effect follows the administration of an inert substance, along with the suggestion that the subject will get worse.

Brain imaging techniques have been crucial to understanding the neurobiology of negative expectations, and most of this research has been performed in the field of pain. Overall, negative expectations may result in the amplification of pain, and several brain regions, like the anterior cingulate cortex, the prefrontal cortex, and the insula, have been found to be activated during the anticipation of pain.

Besides neuroimaging techniques, pharmacological studies give us insights into the biochemistry of the nocebo effect and of negative expectations. A model has recently been proposed whereby the opioidergic and the CCK-ergic systems may be activated by opposite expectations of either analgesia or hyperalgesia, respectively. In other words, verbal suggestions of a positive outcome (pain decrease) activate opioid receptors, while suggestions of a negative outcome (pain increase) activate CCK receptors. The involvement of CCK in nocebo hyperalgesia is likely to be mediated by anxiety, as benzodiazepines have been found to block both nocebo-induced hyperalgesia and the typical anxiety-induced hypothalamus-pituitary-adrenal hyperactivity. Conversely, the CCK antagonist, proglumide, has been found to prevent nocebo hyperalgesia but not the hypothalamus-pituitary-adrenal hyperactivity, which suggests two independent biochemical pathways activated by nocebo suggestions and anxiety.

More recent studies have found that nocebo effects are also associated with a decrease in dopamine and opioid activity in the nucleus accumbens, thus underscoring the role of the reward and motivational circuits in nocebo effects as well. In other words, the activation/deactivation balance of both dopamine and opioids in the nucleus accumbens would account for the modulation of placebo and nocebo responses. Therefore, different neurotransmitters, such as CCK, dopamine, and opioids, are involved in the complex psychosocial modulation of pain perception by placebos and nocebos.

Fabrizio Benedetti

See also Brain Imaging; Context Effects in Perception; Migraine; Mind and Body; Multimodal Interactions: Pain–Touch; Pain: Assessment and Measurement; Pain: Cognitive and Contextual Influences; Top-Down and Bottom-Up Processing

Further Readings

Benedetti, F. (2008). Mechanisms of placebo and placebo-related effects across diseases and treatments. *Annual Review of Pharmacology and Toxicology, 48,* 33–60.

Benedetti, F. (2008). *Placebo effects: Understanding the mechanisms in health and disease.* Oxford, UK: Oxford University Press.

Colloca, L., & Benedetti, F. (2005). Placebos and painkillers: Is mind as real as matter? *Nature Reviews Neuroscience, 6,* 545–552.

Enck, P., Benedetti, F., & Schedlowski, M. (2008). New insights into the placebo and nocebo responses. *Neuron, 59,* 195–206.

Petrovic, P., Kalso, E., Petersson, K. M., & Ingvar, M. (2002). Placebo and opioid analgesia—Imaging a shared neuronal network. *Science, 295,* 1737–1740.

Price, D. D., Finniss, D. G., & Benedetti, F. (2008). A comprehensive review of the placebo effect: Recent advances and current thought. *Annual Review of Psychology, 59,* 565–590.

Scott, D. J., Stohler, C. S., Egnatuk, C. M., Wang, H., Koeppe, R. A., & Zubieta, J. K. (2008). Placebo and nocebo effects are defined by opposite opioid and dopaminergic responses. *Archives of General Psychiatry, 65,* 220–231.

Wager, T. D., Rilling, J. K., Smith, E. E., Sokolik, A., Casey, K. L., Davidson, R. J., et al. (2004). Placebo-induced changes in fMRI in the anticipation and experience of pain. *Science, 303,* 1162–1166.

PAIN: TREATMENTS FOR CHRONIC

Chronic pain is a prevalent health problem that many contend is undertreated in most societies. According to the World Health Organization, the prevalence of chronic pain is between 20 to 30% of the population worldwide, including over 50 million Americans. The most commonly reported chronic pain problems are low back pain, headaches, and joint pain. Chronic pain is defined as pain that has persisted for three or more months. As opposed to acute pain, where the cause of pain is usually evident, the cause of chronic pain in many people is elusive. Even in diseases where an underlying source of pain can be easily identified, such as osteoarthritis, there is often little association between the degree of joint degeneration and the intensity of pain reported by the patient. In fact, many people who report having no pain have evidence of abnormalities often associated with pain conditions, such as disc bulges or even disc herniations in the spine, or evidence of moderate to severe arthritis in a joint. Given that pain is a subjective experience that is influenced by many factors, how pain is experienced plays a significant role in seeking out care, identifying disability, and ultimately determining the societal cost and burden of treating chronic pain. This entry provides only a brief overview of some of the methods used to treat chronic pain.

Underlying Causes of Chronic Pain

Although chronic pain may be the result of misdiagnosed, mismanaged, or persistent acute pain, many believe that the development and maintenance of chronic pain is a complex process that involves the interplay between psychosocial and physical factors. According to this biopsychosocial model of pain, pain is conceptualized as a multidimensional phenomenon that is influenced by many factors, including mood, previous experience, and cultural and other beliefs about pain, in addition to sensory input. It is also recognized that people with pain demonstrate widely different emotional and behavioral responses to similar levels of pain intensity. Thus, how a person copes with or responds to acute pain is an important determinant of the development of chronic pain.

In the transition from acute to chronic pain, the source of nociceptive input may change from the periphery to the central nervous system. This phenomenon, referred to as central sensitization, is thought to occur in response to repeated stimulation from peripheral pain fibers. Constant signaling over time is believed to sensitize pain neurons in the spinal-thalamic tract, leading to enhanced responses to peripheral nociceptive input and spontaneous firing. Thus, pain signals to the brain may be amplified or produced in the spinal cord, even when peripheral pain input is diminished or absent. In addition, premorbid physical factors may also influence the development of chronic pain. There is emerging evidence that genetics play a role in human pain perception. For example, genes that influence catechol-O-methyl transferase (COMT), an enzyme that breaks down catecholamines, GTP cyclohydrolase, a building block in other amino acids, and Nav1.9, a sodium channel that is preferentially expressed in nociceptive neurons in the dorsal root ganglia, have all been found to be associated with human pain sensitivity. In addition, prospective studies have found that genes associated with COMT predict the development of chronic pain conditions such as temporomandibular joint disorder and knee osteoarthritis. These genetic factors may augment the intensity of any acute pain problem, increasing the likelihood of developing chronic pain.

Specific Treatments

Multidisciplinary Treatment

Consistent with the biopsychosocial model, the treatment of chronic pain needs to address the multitude of factors that influence the development and maintenance of the disorder. However, people with chronic pain are highly heterogeneous, and many advocate that patients need to be matched to the most appropriate treatments. Initial studies addressing this issue have suggested that people with chronic pain whose pain is not very intense, or whose pain has little psychosocial or functional impact, tend to benefit equally from simple interventions as well as more intense interventions. In contrast, people with chronic pain who also display significant psychosocial or functional impairment tend to benefit most from interdisciplinary or multidisciplinary approaches. The specific components these programs generally include are the following: (a) medical care, (b) exercise or specific physical therapy intervention, (c) psychological intervention, (d) occupational therapy, and (e) vocational rehabilitation services. Research suggests that multidisciplinary pain treatment for chronic pain is superior to no treatment or more simple interventions, and although these programs can be costly, they have been suggested to be cost-efficient when compared to other interventions for chronic pain.

Often, the goal of multidisciplinary treatment is not to decrease pain, but to teach people how to manage it better or to restore function even in the face of pain. Despite this, studies report that people treated using this approach often report that their pain is reduced. A relatively new multidisciplinary approach to treating chronic pain involves in-vivo exposure to light normal activity. There is growing evidence to suggest that beliefs that pain is a signal of harm or damage to the body and that one should avoid activities that may cause pain play a significant role in the development of disability due to chronic pain, particularly in conditions such as chronic back pain. Treatments focused on restoring movement and function have been found to significantly decrease pain-related disability. Although the majority of the studies on these approaches have been conducted in people with chronic back pain, these approaches appear to be beneficial in treating other chronic pain conditions as well.

Medications

Several types of medications are used to treat chronic pain, despite the fact that there are few well-controlled studies examining the efficacy of various medications for the treatment of chronic pain that is not due to cancer. In addition, there are few standardized guidelines that outline specific strategies for medication management of chronic pain, and some medications tend to be used only for certain disorders. A brief overview of the more commonly administered medications is presented next.

Nonsteroidal anti-inflammatory drugs (NSAIDs) are thought to be useful in conditions where inflammation plays a role in pain. However, NSAIDs also directly influence pain transmission as they stimulate the production of prostaglandins, which in turn desensitize nerves that transmit pain. A side effect of these medications is that they may cause stomach upset and even gastrointestinal bleeding. Newer NSAIDs have been developed that reduce the risk of these gastrointestinal complications, but some of these drugs (called COX-2 inhibitors) have been found to increase the risk of heart attack and stroke and have been withdrawn from the market (e.g., rofecoxib).

Opioids are frequently used to treat chronic pain, but their use is somewhat controversial.

While these medications (such as codeine, morphine, and fentanyl) are potent analgesics, their efficacy for treating chronic nonmalignant pain has not been well demonstrated. Although patient advocacy and professional groups suggest that these medications can be safe and effective in the hands of an experienced pain practitioner, concerns regarding the potential for abuse and addiction frequently arise, as prescription medications are being increasingly misused in the United States. Also, because the treatment of chronic pain is long term, the development of tolerance to opioids over time necessitates increasing the dose of the opioid to achieve adequate analgesia.

Opioids are frequently combined with either acetaminophen or NSAIDs, as these combinations have been found to increase the pain-relieving effect of the opioid. These drugs come in long- and short-acting forms, and drug companies are beginning to prepare formulations that deter abuse. Tramadol is a synthetic opioid that is commonly prescribed in the United States, as studies on this medication in Europe suggested that it was less likely to cause addition or dependence compared to natural opioids. Recently, research has suggested that people taking long-term opioids may develop opioid-induced hyperalgesia. While the exact mechanisms that cause this are not known, studies have shown that some patients with chronic pain display less sensitivity to experimental pain after being withdrawn from opioids compared to when they were taking the medication.

Antidepressant medications can be beneficial in treating chronic pain, and these medications appear to produce analgesia independent of their influence on mood. The antidepressants that appear to be most effective for treating pain are those that exert some influence on both serotonin and norepinephrine. These medications include the serotonergic-noradrenergic reuptake inhibitors such as duloxetine, which is FDA approved for treatment of diabetic neuropathy and post-herpetic neuralgia, and the older tricyclic antidepressants, such as amitriptyline. The mechanisms by which these medications influence pain are not entirely clear, although there is some suggestion that they may influence descending pain inhibitory pathways in the brain, which in turn inhibits pain transmission in the spinal cord and periphery.

Anticonvulsants such as gabapentin and pregabalin have also been shown to be efficacious in treating chronic pain. The mechanism of action of these drugs is not known, although they may act to decrease abnormal cell firing in pain neurons. Although these drugs historically were thought to be of benefit only when treating neuropathic or nerve-related pain, pregabalin is also FDA approved for treating fibromyaglia, a condition characterized by widespread tenderness or pain in the muscles.

Physical Therapy/Exercise

Physical therapy and exercise have been found to be beneficial in treating a number of chronic pain conditions, for various reasons. First, some chronic pain conditions, such as back pain, are believed to be caused by musculoskeletal dysfunctions that restrict the mobility of the spine or other joints, or alternatively cause them to be too mobile. Thus, treatment and exercises that restore joint mobility, or reduce hypermobility, can serve to decrease pain. Second, general exercise may be an important component in the treatment of chronic pain, as inactivity due to pain may lead to muscle weakening and deconditioning. Improving strength, endurance, and mobility may help to decrease disability. Finally, exercise may alter neuropeptides, which in turn may influence pain perception. For example, bouts of exercise are known to subsequently decrease the perception of experimental pain, a phenomenon referred to in the literature as exercise-induced analgesia.

Psychological Treatment

Cognitive-behavior therapy (CBT) for chronic pain involves working with a psychologist to learn different cognitive and behavioral strategies for managing pain, stress, and mood disturbance. While this treatment is typically offered as part of a package, and the components may vary across providers, CBT usually involves the following: (a) altering maladaptive beliefs about pain, (b) working with patients to expand the repertoire of ways in which they can cope with pain, (c) relaxation training, (d) improving assertiveness and communication skills, and (e) education about the nature of pain. Reviews of CBT suggest that this intervention is beneficial in improving pain, function, and mood disturbance among people with various chronic pain conditions.

Biofeedback is a technique that is used to provide a patient with chronic pain information about a particular bodily state. This treatment is based on the premise that making people more aware of a particular physiological signal can help them gain greater control over that bodily function. The types of biofeedback employed tend to vary based on the particular pain disorder. For example, in treating chronic back pain, many believe the disorder is associated with abnormalities in muscle tension or muscle firing patterns during movement. Thus, biofeedback for back pain typically entails providing the patient with feedback on muscle tension using a surface electromyography monitor, and teaching people to lower their tension, or normalize their muscle activity. However, few studies have systematically examined the efficacy of biofeedback interventions for treating chronic pain.

Injections/Interventional Procedures

The goal of injection or interventional therapy is to block the source of pain. The success of these procedures relies upon finding a specific "pain generator" in the area of the body where one experiences pain. There are many types of injections, but most involve administering steroids to reduce inflammation and/or the use of anesthetics. These procedures do not necessarily provide lasting relief, but can provide diagnostic information about possible sources of pain that can then be treated with methods that provide longer lasting relief, such as radiofrequency ablation. To date, few studies have examined the efficacy of these procedures for various chronic pain problems.

Implantable Devices

Spinal cord stimulation can be used to treat pain that is confined to a single branch of nerves that comes out of the spinal cord. An electrode is implanted near the nerve root that generates an electrical impulse that masks or overrides the person's chronic pain. Intrathecal pumps are also used to treat chronic pain. These devices continuously deliver medication, such as morphine, directly into the spinal cord. There is evidence that this method of delivering medication may be more effective than taking medication orally, and it may also reduce the side effects.

Michael E. Geisser

See also Cutaneous Perception; Pain: Cognitive and Contextual Influences; Pain: Neuromatrix Theory; Pain: Physiological Mechanisms

Further Readings

Loeser, J. D., Butler, S. H., Chapman, C. R., & Turk, D. C. (2000). *Bonica's management of pain.* Philadelphia: Lippincott, Williams and Wilkins.

Melzack, R., & Wall, P. (Eds.). (2003). *Handbook of pain management* (5th ed.). Philadelphia: Elsevier/ Churchill Livingston.

PERCEPTION IN UNUSUAL ENVIRONMENTS

Humans evolved to walk on land and deal with nearby objects. They later used boats, rode horses, and traveled in carriages. The development of engines led to high-speed travel in cars, trains, and aircraft. Humans also started diving underwater using various gas mixtures. They explored most of the world and landed on the moon. These activities stretched human physiology to the limits, and depended on artificial habitats or survival apparatus. The environments also stretched human perceptual systems, which had not evolved to cope with all possibilities. Perceptual adaptation to some sensory transformations is possible, but we cannot replace total losses of sensory information. This entry describes how various environments affect visual, auditory, and vestibular perceptions.

Perception at a Distance

Early humans lived as hunter-gatherers, and their perceptual and motor skills were adequate for the near environment. When humans travel further or climb mountains, problems arise because the distant spatial information is inadequate or misleading. People make mistakes about the size and distance of objects, the height and steepness of mountains, and whether the ground below is flat or rising.

Spatial Judgments

Distance judgments are affected by many factors, such as atmospheric visibility, the presence of intervening objects, and the direction of viewing. With horizontal viewing, observers increasingly underestimate far distances, particularly across featureless terrain. When looking up or down from a height, distances usually appear large and objects small. A well-known example is the moon illusion: The moon appears large and near on the horizon, but small and distant when up high. Slope judgments are also distorted, probably through distance foreshortening. When viewed from a height, distant downhill slopes appear too flat, a flat valley appears to rise, and facing uphill slopes appear too steep. The perceived horizontal is attracted toward the angle of the nearest frontal slope, affecting the perceived height of further hills. Observers looking over a gentle downhill slope judge horizontal eye level as too low and the opposite hilltop as too high; and vice versa when looking up a slope.

Atmospheric and Weather Effects

We are accustomed to the usual atmospheric visibility where we live. One of the cues used in judging distance is aerial perspective—the reduction of luminance and color contrast with distance. The light reflected from objects becomes more similar to that of the sky background as it travels through the atmosphere to reach the observer. In a misty atmosphere we overestimate distance, and in a clear atmosphere we underestimate it. Mist is particularly dangerous while driving, because distant objects are invisible or appear far away. An exceptionally clear atmosphere can be misleading for walkers, who think that distant land is close, and attempt to walk impossible journeys. They may also see baffling mirages.

Walking or skiing in a whiteout is hazardous. Walkers can lose their sense of orientation, and skiers can lose their sense of their own movement. If stationary skiers are surrounded by moving skiers, they may feel that they are moving in the opposite direction, apply an inappropriate correction and fall over in the same direction as the moving skiers.

Underwater

Optical Distortion and Loss

The human eye evolved to work in air and becomes farsighted in water. The corneal surface loses its refractive power because it has a similar refractive index to water. Therefore, divers wear a facemask, allowing the eyes to operate in air. However, the difference in refractive index between air and water causes refraction of light rays at the interface of the facemask. The image size is enlarged by about 4/3, and the optical distance of the image is about 3/4 of the object's physical distance. This spatial distortion causes problems in the perception of size, distance, speed, slope, curvature, and location. Divers misreach at first, but partially adapt to the distortions and show brief aftereffects in air. Perceptual *trading* may occur between incompatible distortions, such as left and right angular displacements—divers adapt to one aspect but make larger errors on the other. Divers partially adapt to both distance and size: Objects usually appear to be located further than their optical distance, but enlarged by less than 4/3. Apparent distance is affected by the clarity (density of particles) of the water; objects appear further away in murky than in clear water. This is similar to the effects of aerial perspective, but water has much lower clarity than air. Consequently, distance is increasingly overestimated with viewing distance in water, whereas in clear air it is increasingly underestimated.

It might be thought that the enlarged image size (magnified by 4/3 through refraction) would improve visual acuity under water compared to the same test in air. Unfortunately, any gain is offset by the lack of water clarity and reduced illumination with depth: Visual acuity, contrast sensitivity, and stereoscopic acuity are poorer than in air. The shape of the facemask also cuts off some peripheral vision, forcing divers to move their heads when searching the visual scene.

There is also a distortion and reduction of color vision, the effects varying with the depth of the water and the viewing distance. Pure sea water has a blue cast because the blue wavelengths are scattered and the red are absorbed. The presence of yellowish vegetable material makes coastal water green and inland water yellow or red. When viewed horizontally, objects take on the predominant color of the water, the effect increasing with viewing distance. Some perceptual adaptation occurs to the predominant color, as it does to changes of illumination in air, but divers nevertheless make mistakes about the absolute color of objects. When the diver descends, there is a more serious loss of color vision. The red wavelengths are progressively absorbed with depth, and red objects appear first green and then black. The brain cannot compensate for the total loss of a wavelength, and the use of an artificial light source is the only solution.

Auditory Distortion and Loss

The human ear functions poorly when the auditory canal is filled with water, causing some loss of the ability to detect quiet sounds. There is also a loss of directional sensitivity, which is about 10° in water and 1° in air. This loss is partly due to the absence of an intensity difference at the two ears, the head casting no sound shadow in water. In addition, the difference in time of arrival at the two ears is reduced because sounds travel about four times faster in water than in air. This makes the difference hard to detect and distorts the apparent location toward the median plane. All sounds appear to come from an auditory area of 34° at the front or back, sounds at 90° in water being equivalent to 17° front or back in air. Divers can partially overcome the ambiguity of the sound source by moving their heads to and fro.

Orientation and Motor Skills

The perception of bodily orientation is impaired under water. Buoyancy reduces information from touch and from the joint receptors, and the visual scene may offer few guides to the vertical. However, the vestibular balance system of the inner ear operates normally, apart from the rotary vertigo that some divers suffer because of difficulties in equalizing the pressure in the middle ear. Motor skills are impaired by optical distortion, buoyancy, and cold or gloved hands. The weight of objects is reduced in water, owing to the upthrust of the water, but divers show some degree of mass constancy or weight adaptation. Their discrimination between masses is also reduced. Finally, all performance can be impaired at depth through anxiety and sometimes nitrogen narcosis (a toxic effect of breathing nitrogen at pressure).

Aerospace and Travel

Altered Force Environments

Humans evolved to withstand the 1 *g* accelerative force of earth's gravity, which acts through the vertical axis of an upright body. Passengers experience greater linear accelerations at the start or stop of vehicular travel in a straight line, and angular accelerations at the start or stop of a turn to the left or right. High *g* forces can be produced in the human centrifuge, and fairground rides and sea travel also cause a variety of *g* forces. Any type of unusual travel induces false perception of movement and orientation, and sometimes motion sickness. The normal correspondence between visual, vestibular, and tactile stimulation breaks down, and travelers feel disoriented. Some travelers adapt quickly, but it can take two or three days at sea or in space to suppress sickness, and longer for full perceptual adaptation.

Weightlessness

Weightlessness can be achieved in various ways. Brief periods of up to 25 seconds of microgravity (near 0 *g*) can be produced in parabolic flight, preceded and followed by periods of about 2 *g*. Experiments can be carried out during these flights, but participants have no opportunity to adapt to long-term microgravity. Longer spaceflights are needed for that. An orbiting spacecraft is in a state of microgravity because the earth's gravitational acceleration is exactly balanced by the radial acceleration produced by the spacecraft's curved flight path. Microgravity causes unusual sensory conflict because the semicircular canals of the vestibular system operate normally, but the otolithic organs (little stones that slide over hair cells) do not. The canals respond mainly to rotary accelerations, and the otoliths to linear accelerations. Under microgravity, the otoliths can no longer indicate head orientation with respect to gravity, but only linear acceleration forward-backward, left-right, or up-down. Therefore, astronauts must learn to reinterpret otolithic information. Pitch and roll head movements are particularly provocative of motion sickness, as are ambiguous visual stimuli, such as the view of the spacecraft from an "inverted" orientation, or seeing another crew member "upside down." Unlike terrestrial travelers, astronauts can remove themselves from provocative stimulation: They can wedge themselves into a corner, close their eyes, and keep their head still.

Motor Skills in Varied Force Environments

People easily adapt their perceptual-motor skills to altered *g* forces. They learn new patterns of hand-eye coordination and reprogram rapid arm movements so as not to overreach in low *g* or underreach in high *g*. They partially adapt to their own changes in arm weight and learn to judge the mass of objects through inertial information (by moving them), rather than from static weight information. However, mass discrimination under microgravity remains poorer than weight-plus-mass discrimination on earth because of the loss of continuous weight information.

Aftereffects

Divers have brief aftereffects after a short stay underwater, but returning sailors and astronauts take a few days to recover their "land legs" or "earth legs." Astronauts feel heavy and clumsy. Their vestibular coordination is disturbed, and they may see the world swing around with head movements. Some aftereffects may be due to the continued reinterpretation of otolith signals as linear accelerations rather than as head tilt, and there may also be changes in sensitivity to linear acceleration in different body axes. Other aftereffects are nonvestibular, such as feelings of heaviness, impairment of weight discrimination, underreaching for objects, and faulty awareness of limb position: They may be due to changes in proprioceptive and tactile sensations.

Helen E. Ross

See also Atmospheric Phenomena; Auditory Localization: Psychophysics; Color Perception; Constancy; Perceptual-Motor Integration; Self-Motion Perception; Spatial Layout Perception: Psychophysical; Vestibular System; Weight Perception

Further Readings

Adolfson, J. A., & Berghage, T. E. (1974). *Perception and performance under water.* New York: Wiley.

Lackner, J. R., & DiZio, P. (2000). Human orientation and movement control in weightless and artificial gravity environments. *Experimental Brain Research, 130,* 2–26.

Minnaert, M. (1954). *The nature of light and colour in the open air* (H. M. Kremmer-Priest, Trans., & K. E. Brian Jay, revision by). New York: Dover.

O'Shea, R. P., & Ross, H. E. (2007). Judgments of visually perceived eye level (VPEL) in outdoor scenes: Effects of slope and height. *Perception, 36,* 1168–1178.

Reschke, M. F., Bloomberg, J. J., Harm, D. L., Paloski, W. H., Layne, C., & McDonald, V. (1998). Posture, locomotion, spatial orientation, and motion sickness as a function of space flight. *Brain Research Reviews, 28,* 102–117.

Ross, H. E. (1974). *Behaviour and perception in strange environments.* London: Allen and Unwin.

Ross, H. E. (1989). Perceptual and motor skills of divers under water. *International Reviews of Ergonomics, 2,* 155–181.

PERCEPTUAL DEVELOPMENT: ATTENTION

Attention is a broad term that refers to a series of complex processes. Although attention influences the function of all sensory systems, this entry will consider the development of visual attention because most research considering the development of attention has focused on the visual system.

Developmental psychologists have often taken paradigms that have been used to study attention in the primate visual system and adapted them to function as "marker tasks" for the development of attentional networks in humans. Three attentional systems have been identified. These include (1) arousal, (2) orienting, and (3) executive control. This entry defines these systems and considers their development in turn.

Arousal is a sense of heightened engagement. For example, we become alert or aroused into action when a door suddenly slams. Orienting refers to either shifting the eyes to some location in space (overt orienting) or shifting attentional focus to some location in space without necessarily looking there (covert orienting). Finally, executive control is the ability to suppress automatic responses and exercise control over thoughts or actions. For example, one may automatically look to the left while crossing the street, even as it is a one-way street with traffic coming from the other direction.

Development of each of these abilities plays an important role in infants' perceptual development. There is a benefit to the emergence of skills that allow infants to make attention-directed eye movements to explore their environment, to become more capable of meaningfully sustaining attention, and to control their own actions. They become active participants in their own perceptual processes. The three to six month age range turns out to be an important time for the emergence of attentional processes. Notably, many of these processes continue to develop well into childhood and, in some cases, into adolescence.

Development of Attentional Processes

Arousal refers to the engagement and maintenance of attention in the service of information acquisition, cognition, or behavior. Psychophysiological measures, specifically changes in heart rate, have been combined with simple viewing paradigms to examine changes in arousal and sustained attention. Heart rate changes during sustained attention and attention termination have provided researchers with important information about the early postnatal development of the arousal system. There is generally a slowing of heart rate during the attentive or vigilant state. This returns to baseline levels at attention termination (i.e., inattentiveness even as the subject continues to look at the stimulus). This highlights the idea that attention is more than vision. Many of us can recall a situation in which we are watching a television show but cannot recall anything that happened. The extent of arousal is reflected by the extent of heart rate deceleration during sustained attention. Between three and six months of age, infants show greater heart rate deceleration during sustained attention and simultaneously behavioral indication of better performance on a variety of measures (e.g., more effective information acquisition).

The orienting system has also been heavily studied in infancy. For example, infants as young as two months of age can learn to direct eye movements to anticipate the appearance of a picture in a particular location in space. This requires shifting focus to a location in the absence of something to look at. By definition, this is beyond the scope of simple alerting. Spatial cueing paradigms have been used most extensively to examine the development of overt and covert orienting. In general, attention is engaged with a centrally presented attractive stimulus. A brief peripheral "cue" stimulus is flashed on a screen positioned in front of the infant, to the right or to the left of center. If an eye movement or saccade is allowed to the peripheral cue, the paradigm is considered to index overt orienting. If the cue is too brief to elicit an eye movement, the paradigm is measuring covert orienting. After a delay interval, the peripheral cue is presented by itself either in the same location (valid trial) as the flash or in the opposite (invalid trial) screen location. Changes in the speed of an eye movement, or the frequency of looking in very young infants, to the target as a function of a previous cue are taken as indication that an attentional shift was made to the flashing peripheral cue. Eye movements to the target will be faster, or more frequent, on valid trials relative to invalid trials when the delay between presentations is short. That is, responses are *facilitated* by the appearance of the peripheral cue. However, they will be slower, or relatively infrequent, if the delay between cue and target presentations is long. This effect has been called inhibition of return (IOR). In the case of overt spatial cueing, IOR has been reported in newborn infants. Inhibition of return following covert shifts of attention may begin to emerge at around three months of age but is not observed consistently until six months. The IOR effect is considered to play a role when infants search the visual environment, prohibiting eye movements to previously examined locations. This would theoretically have the effect of maximizing the number of locations targeted in a visual scene. Indeed, IOR is a putative mechanism at play in tasks that are specifically designed to examine how infants scan the visual world.

More information exists in a natural visual scene than can be processed by the system at any one time. The natural environment, cluttered with many objects of varying colors, textures, and movements, consistently presents us with this problem. In a visual search task, one target element is placed in an array of distracters. Search difficulty increases with target/distracter similarity. A famous example of effortful visual search would be a "Where's Waldo?" display. Successful search requires the viewer to resolve the competition between similar elements to determine the next place to orient attention. This process has been called visual selective attention (i.e., the ability to select relevant information in the presence of competing distracters). Visual search paradigms have been applied to examine orienting under complex viewing conditions and have indicated that infants provide evidence of sensitivity to the competition induced by complex visual displays as early as three months of age. That is, they will show a pattern of increasingly longer search times as displays become more complicated. For example, it takes increasingly longer time to find Waldo in a "Where's Waldo?" display as extra figures are added to the scene.

Studies examining the development of executive control, the ability to overcome automatic responses and control behavior, suggest that skills emerge in infancy but continue to develop through adolescence. One paradigm that shows such development is the antisaccade task. Subjects fixate on a centrally presented stimulus as a peripheral target is flashed to the right or left of center. Until this point, this and the spatial cueing paradigm are quite similar. However, with antisaccade tasks, subjects are required to make an eye movement to the side of the screen opposite to where the peripheral flash occurred. This requires suppression of the automatic eye movement to the location where the target occurred and maintenance of the rule to look to the opposite location. Although research indicates that infants as young as four months of age provide evidence of some effective performance on antisaccade tasks, research shows continued improvement on this task in childhood and through adolescence.

Implicated Neural Circuitry

The neural systems supporting attentional functions have been studied extensively using both primate models of the mammalian visual system and functional magnetic resonance imaging (fMRI) in human children, adolescents, and adults. Both allow for precise recordings from particular brain regions as a subject performs a task, and both can inform which circuitry is disrupted in the case of a subject with attentional problems. This type of research has often used many of the same paradigms used with infants, making it possible for infant attention researchers to suggest that behavioral development in these paradigms serves as a marker for change in neural pathways identified in the human and primate work.

In short, findings indicate that newborns target eye movements in a reflexive fashion, driven by external salient information, and largely recruit a pathway involving the retina, the lateral geniculate nucleus, the primary visual cortex, and the superior colliculus. The emergence of attentional control over orienting may stem from the development of certain cortical regions, including pathways further involving the visual cortex and parts of the parietal and prefrontal cortices. The parietal cortex is engaged in covert shifts of attention. The prefrontal cortex plays a role in voluntary control over eye movements involving delays, presumably as a function of the requirement to maintain information over the delay period. Recent work using event-related potentials and source localization techniques in conjunction with the covert spatial cueing procedure found greater spatial cueing modulation of the extrastriate cortex, located next to the primary visual cortex, in five-month-old relative to three-month-old infants, presumably as a function of increased cortical control over this visual area.

The Role of Attention in Perception

The emergence of attentional skills must have some impact on perception. Each opens up a different window into exploration of the external environment, which in turn influences perception. To test this idea empirically requires asking whether changes in attention in a group of infants occur simultaneously with changes in some percept. For example, *perceptual completion* refers to the ability to perceive disjoint object parts as complete. For example, take a large pen and place it under an envelope on a desk. Visually, we see only the top and bottom parts of the pen, as the envelope occludes the middle. Despite this input, our visual system completes this object and perceives it as one pen, and not as two separate parts. Perceptual completion, which is important for effective visually guided navigation and interaction with the environment, develops over the first several postnatal months. Studies using experimental displays similar to the pen/envelope situation (a moving bar hidden behind a box) have shown that *where* on the display three-month-old infants look is very important to whether they provide evidence of perceiving the object as unified or broken. For example, those who look more at the top and bottom of the bar, relative to other display regions, are more likely to indicate perceiving the object as a unified bar, as adults do. What role does the development of attentional orienting play in this object perception? To address this question, three-month-old infants were observed for evidence of unity perception in a partly occluded object display and tested as well in a visual search paradigm. In the visual search task, experimenters manipulated the competition between display elements (one target and multiple distracters) in a visual scene. Infants who provided evidence of perceiving the unity of the bar parts showed efficient scanning patterns during exploration of the display. Importantly, these same infants also provided evidence of visual search behavior indicative of voluntary attention-guided orienting.

This is only one example of a study that attempts to precisely define a relationship between the development of attentional capacities and perception. Overall, a consensus exists that perception is active. Studying the development of attention brings us closer to understanding how visual exploration strategies change over early postnatal development. This, in turn, has broad implications for how effectively information is acquired from a cluttered environment for subsequent perception and learning.

Dima Amso

See also Action and Vision; Attention: Covert; Attention: Effect on Perception; Attention: Selective; Infant Perception

Further Readings

Amso, D., & Johnson, S. P. (2006). Learning by selection: Visual search and object perception in young infants. *Developmental Psychology, 42,* 1236–1245.

Dannemiller, J. L. (1998). A competition model of exogenous orienting in 3.5-month-old infants. *Journal of Experimental Child Psychology, 68,* 169–201.

Johnson, M. H. (1995). The inhibition of automatic saccades in early infancy. *Developmental Psychobiology, 28*(5), 281–291.

Johnson, M. H. (2005). *Developmental cognitive neuroscience* (2nd ed.). London: Blackwell.

Posner M. I. (Ed.). (2004). *Cognitive neuroscience of attention.* New York: Guilford.

Richards, J. E. (2005). Localizing cortical sources of event-related potentials in infants' covert orienting. *Developmental Science, 8*(3), 255–278.

PERCEPTUAL DEVELOPMENT: COLOR AND CONTRAST

Spatial contrast is arguably the most fundamental variable in vision. It enables us, among other things, to discriminate figure from ground, to find the edges of objects so that those objects can be recognized by their shapes, to match regions in the two eyes to enable depth perception, and to segregate regions in the visual image that differ in texture. Without the ability to perceive contrast, these fundamental aspects of visual perception would be all but impossible. This entry focuses on the development of spatial contrast vision and color vision.

One can also make the argument that color plays an important role in many of the same aspects of visual perception. Regions in the visual image that have the same texture but different colors can be readily distinguished by a perceiver with normal color vision. Those regions probably differ in their material properties. Color also plays a role in object recognition and discrimination. Ripe bananas look very much like unripe bananas in their shapes, but they differ considerably in their colors. Even in the absence of luminance contrast, an organism with color vision can find edges in the visual image between regions that differ only in their hues.

Given the importance of color and contrast to vision in general, it is logical to ask about the development of color and spatial contrast mechanisms. Contrast here refers to spatial differences in luminance. In adults, sensitivity to contrast is measured by finding the smallest difference in luminance between adjacent light and dark bars that can be reliably detected. At their most sensitive, adults can detect a difference in the luminances of adjacent bars of as little as 0.1% (the inverse of this is taken as contrast sensitivity; in this case 1/.1% = 1000). The sizes of these bars can be varied to determine how contrast sensitivity varies with spatial scale. Spatial scale is typically quantified by the number of adjacent light and dark bars that fit within some distance, such as a degree of visual angle (approximately the distance covered by the width of your thumb viewed at arm's length). When quantified in this way, spatial scale is referred to as spatial frequency with units of cycles per degree (cpd). A pair of adjacent light/dark bars comprises one cycle, so spatial frequency refers to the number of such cycles in a degree of visual angle. High spatial frequencies refer to gratings with many of these light/dark cycles in a degree of visual angle. Adults with normal vision can see gratings that have as many as 60 of these light/dark cycles in a degree of visual angle (60 cycles per degree). When the spatial frequency exceeds 60 cpd, adults can no longer see such fine detail. This highest detectable spatial frequency is referred to as acuity.

The Development of Spatial Contrast Vision

Human infants come into the world with very poor contrast sensitivity relative to adults. For example, it can take between 25 and 100 times as much contrast for a 5-week-old to see a grating as for an adult to see the same grating. Poor contrast sensitivity reduces the amount of information available in the visual image for subsequent processes, such as figure-ground segregation and object recognition. The "window on the world" provided by contrast is much smaller for the human neonate than it is for adults or even for older infants and children. As an example, consider a very simple spatial pattern, a square-wave grating: alternating dark and light bars with sharp edges at the transition from dark to light and vice versa. Seeing the edges of these bars as sharp requires good contrast sensitivity. A perceiver with poor contrast sensitivity would see a square-wave grating as a pattern with "fuzzy" edges between the light and dark bars—as if the sharp edges of the square wave pattern had been blurred or smeared. Human neonates and young infants are just such perceivers. The information available to important perceptual processes, such as figure-ground segregation and object recognition, is severely diminished for human infants.

Contrast sensitivity develops rapidly over the first year of life and beyond. At 7.5 months of age, infants show contrast sensitivities at some spatial frequencies that are essentially equivalent to those shown by adults when measured using the visual evoked potential (VEP). The VEP records electrical activity at the scalp that reflects the responses of many underlying neurons, in this case in response to a flashed grating. Behavioral measures show that there is continued development of contrast sensitivity until at least 7 years of age. There are several factors that contribute to this development. One might expect changes in optical quality and accommodation—the ability to focus at the appropriate distance—to play a role in the development of contrast sensitivity, but they play only minor roles relative to other factors. The primary reason for the poor contrast sensitivity is that the retina itself is quite immature at birth. Especially in the fovea, the cone photoreceptors at birth are packed much less densely than they are in adults. One consequence of this is that many fewer photons are absorbed in the neonate's retina than in the adult's retina given the same pattern of incident light. The photon catch in the neonatal retina is approximately 1/350 as efficient as it is in adults. Spatial contrast sensitivity involves the ability to discriminate neighboring differences in the photon catch over some short period of time (the temporal integration period). Therefore, it follows that a visual system that is working with far fewer photons to resolve these neighboring spatial differences in photon catch rates will be at a disadvantage because the signal with which it is computing these differences will be inherently noisier.

Rapid maturation of the retina is also responsible for the increases in contrast sensitivity and acuity observed postnatally. The photoreceptors begin to pack more densely in the fovea—the central part of the retina where the ability to see fine detail is at its highest. Additionally, the morphology of the cones changes. In particular, the inner segment elongates substantially postnatally. One consequence of this elongation is that photons have a greater chance of being captured and generating a signal than they do when the inner segments are short and wide, as they are in the neonatal retina. Long, narrow inner segments act as wave guides for incident photons channeling the photons to the outer segments where they can be absorbed by a photopigment molecule. Short inner segments, like those that exist in human neonates, are very inefficient wave guides for incident photons. Hence, a large proportion of the incident photons are never absorbed by the neonate's cones.

Although it is true that retinal maturation is responsible for a good deal of the increase in contrast sensitivity during early infancy, more central, cortical mechanisms also play a role. All physical systems, whether biological or fabricated (e.g., an electrical circuit board), have internal noise—random signals that perturb or interfere with processing. Estimates show that the noise in the 7-week-old's visual system is approximately 106 times as great as it is in the adult visual system. Because contrast detection like all detection tasks is limited by noise, it is not surprising that contrast sensitivities are much worse in the human infant than in the adult, given the large difference in intrinsic neural noise. Reductions in this intrinsic noise are also likely to play a role in the improvements of contrast sensitivity during infancy.

The Development of Color Vision

Adults with normal color vision have three kinds of cone photoreceptors responsible for seeing color; for this reason, normal color vision is designated as being trichromatic. Infants are most likely trichromatic by three months of age. Color vision is far from mature at three months, but it does appear that infants possess three functioning cone classes and the circuits necessary to compare the absorptions in those three signal classes. In contrast, infants two months of age and younger routinely fail one or more of the color discriminations that would indicate trichromacy. Researchers do know that the three cone classes are present prior to three months, so these failures of chromatic discrimination probably indicate immaturities in (a) the efficiency with which photons are absorbed, and (b) in the chromatic opponency indicative of vision in color-normal adults.

Trichromacy can be determined from specific wavelength discriminations. In the region above 550 nanometers (nm), color-normal adults are actually dichromatic because the short wavelength sensitive (SWS) cones are too insensitive in this region to provide useful signals. Only medium and long wavelength sensitive (MWS and LWS) cones provide functional vision above 550 nm. The ability to discriminate two different wavelengths in the region above 550 nm must indicate that two cone classes are operative, as well as some mechanism for comparing (i.e., differencing) their signals. Discrimination of two wavelengths in the region above 550 nm is referred to as a Rayleigh discrimination. Similarly, to test for the operation of SWS cones, wavelength pairs (tritan pairs) can be chosen such that these wavelengths are indistinguishable to the MWS and LWS cones. If a pair of wavelengths is indiscriminable to the MWS and LWS cones, but a perceiver can still discriminate them, it must mean that they have another class of cones (SWS). Thus, the ability to make Rayleigh discriminations and tritan pair discriminations can tell researchers whether all three cone classes are present and functioning. Most one-month-olds fail the Rayleigh and tritan discriminations, most three-month-olds can make these discriminations, and two-month-olds are intermediate in their abilities. There are notable discrimination failures, particularly in the short wavelength region, in newborns and neonates. Similarly, most two-month-olds make tritan pair discriminations and most one-month-olds fail them. It should be noted that the reduced visual efficiency (proportion of photons absorbed to photons incident on the retina) noted earlier in the context of contrast vision could also help to explain the pattern of successes and failures in chromatic discrimination early in life.

Adults with normal color vision possess a luminance contrast channel, a red-green (R-G) opponent color channel and a blue-yellow (B-Y) opponent color channel. These channels are thought to be constructed by combining the outputs of the three cone types. For example, the R-G channel is probably constructed by having these cone types converge on neurons that effectively are sensitive to the differences in the responses of the LWS and MWS cones. Similarly, the B-Y channel probably reflects the activity of neurons that are sensitive to the difference between the responses of the SWS cones and the sum of the responses of the

MWS and LWS cones. Chromatic adaptation has revealed R-G opponency in three-month-olds, and there is evidence that the achromatic or luminance channel in infants is probably similar in its characteristics to the same channel in adults. Despite this rapid development of color vision over the first three months, one should not lose sight of the fact that even young children differ from adults on color vision tests, although it is difficult to separate general attentional factors, especially at young ages, from true differences in color vision.

Discriminating color or wavelength does not necessarily tell us how colors appear to infants and young children. Being able to see the difference between two colors does not tell us how those colors actually look to infants and children. We can turn to studies of hue categorization and color constancy to examine the issue of color appearance. Marc Bornstein and colleagues used habituation to determine whether infants perceptually divide the visible spectrum into four hue categories based on three hue boundaries. The habituation method involves repeatedly showing an infant a visual pattern or color in this case until his or her attention wanes, and then showing the infant a new pattern or color. If attention picks up to the new color or pattern, the experimenter makes the inference that the two patterns or colors are discriminable. The four hue categories for adults are blue, green, yellow, and red. Although there were some inconsistencies in the red-yellow region, for the most part, the results showed that four-month-old infants appear to categorize hue in approximately the same way as adults. These earlier studies have been supplemented by later studies using different and probably more sensitive techniques, and the conclusions remain the same in terms of categorical hue perception in infants. The categorical perception of color refers to the fact that although the color spectrum is continuous, adults tend to see it as being divided into qualitatively different regions or categories. Additionally, adults find it easier to discriminate colors that cross these hue boundaries than to discriminate colors that fall within the same category, even if the wavelength difference is the same in both cases.

Interestingly, this tendency to see color categorically is more strongly left-hemisphere-lateralized in adults, but more strongly right-hemisphere-lateralized in four to six month old infants. This probably reflects the impact of language on color perception in adults; this specialization of the left hemisphere for language in adults is not yet present in preverbal infants. Adults show preferences for certain colors over others. Infants also show spontaneous preferences for certain colors over others, and these preferences are similar to adult pleasantness ratings of the same colors. These color preferences shown by infants are probably based on hue or saturation differences rather than on brightness differences across wavelength.

Another aspect of color vision is color constancy. Color constancy refers to the tendency (far from perfect) to discount the effects of the illuminant in perceiving the color of a material or surface. When the same object is seen under different types of lighting, there is a tendency to see the color of that object as being the same despite the fact that the different types of lighting are very different in terms of their color spectra. The light that actually gets to the eye is a product of the light from the illuminant (illuminant spectrum) and the way in which the object reflects that light (reflectance spectrum). The reflectance spectrum is a fixed property of the object, but the light that falls on that object can obviously differ. For example, the spectrum of light from the sun as it is filtered through the atmosphere changes significantly over the course of a day. Despite these large differences in the light that is actually reflected to the eye, adults typically see the color of an object as being nearly the same. There is evidence that rudimentary color constancy is in place by five months of age. Five-month-old infants treat a real or simulated change in the reflectance spectra of simple patterns differently than they do a real or simulated change in the illuminant spectrum. In contrast, two-month-olds do not recognize the same surface reflectance when it appears under a new illuminant. Finally, it is also worth noting that four-month-olds exhibit lightness constancy—a tendency to perceive the reflectance properties of a surface (how black, gray, or white it is), rather than the absolute amount of light being reflected from that surface.

James L. Dannemiller

See also Color Constancy; Color Perception; Color Perception: Physiological; Contrast Perception

Further Readings

Abramov, I., Gordon, J., Hendrickson, A., Hainline, L., Dobson, V., & LaBossiere, E. (1982). The retina of the newborn human infant. *Science, 217,* 265–267.

Banks, M., & Bennett, P. (1988). Optical and photoreceptor immaturities limit the spatial and chromatic vision of human neonates. *Journal of the Optical Society of America A, 5,* 2059–2079.

Bornstein, M., Kessen, W., & Weiskopf, S. (1976). Color vision and hue categorization in young human infants. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 115–129.

Chien, S. H. L., Bronson-Castain, K., Palmer, J., & Teller, D. Y. (2006). Lightness constancy in 4-month-old infants. *Vision Research, 46,* 2139–2148.

Dannemiller, J. (1998). Color constancy and color vision during infancy: Methodological and empirical issues. In V. Walsh & J. Kulikowski (Eds.), *Perceptual constancies: Why things look as they do* (pp. 229–261). Cambridge, UK: Cambridge University Press.

Dannemiller, J., & Banks, M. (1983). The development of light adaptation in human infants. *Vision Research, 23,* 599–609.

Franklin, A., & Davies, I. (2004). New evidence for infant colour categories. *British Journal of Developmental Psychology, 22,* 349–377.

Maurer, D., & Lewis, T. (2001). Visual acuity and spatial contrast sensitivity: Normal development and underlying mechanisms. In C. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 237–252). Cambridge: MIT Press.

Watson, A., Barlow, H., & Robson, J. (1983). What does the eye see best? *Nature, 302,* 419–422.

PERCEPTUAL DEVELOPMENT: FACE PERCEPTION

Faces as perceptual stimuli pack a double punch in terms of being (1) the most extensively experienced class of stimuli that a human observer will encounter over the course of a lifetime, and (2) unique in the sense of conveying a wealth of information (e.g., emotion, gender, race, trustworthiness) that is absent in other objects, even those objects in which we are expert in recognizing. Investigators interested in the development of face perception have been examining the nature of the face representation that infants bring to the task of learning about faces and how that representation changes as a function of differential experience with mother versus stranger, male versus female, same- versus other-race, same- versus other-species, attractive versus unattractive, and positive versus negative emotion. This entry describes newborn face perception abilities in the first year of life and beyond infancy.

Newborn Face Perception Abilities

Newborn infants, just a few minutes from birth, will track with their eyes a schematic visual stimulus resembling a face more than they will track a stimulus that has the external shape of the head but has the internal features of the face scrambled. This result supports the idea that newborn infants enter the world with an internal representation of a face, although some have suggested that the information in the representation may be relatively coarse, consisting of three high-contrast blobs in the correct relative locations for the eyes and the mouth, framed in the contour of a head shape. The coarseness of the representation has actually led to the notion that the newborn representation is not necessarily specialized for faces, but that it reflects a preference for more general perceptual properties (such as symmetry, top heaviness, and congruence) that may also be present in nonface objects. However, the possibility that the initial face representation may be more elaborate is suggested by the finding that newborn infants will imitate facial gestures (such as mouth opening and tongue protrusion) that they see an adult modeling. Whether derived from specialized or more general processes, the newborn representation is believed to equip infants with a mechanism that biases visual attention to the face information present in a visual display. This mechanism may be viewed as adaptive in terms of allowing infants to attend to and recognize members of their own species and also specific people, such as the primary caregiver.

Development of Face Perception in the First Year

If infants have an initial representation of a face as a set of features in a particular arrangement, then the question arises as to how sensitivity develops to the individual features versus sensitivity to the structural whole that incorporates the spatial relations among the features. In addition, for structural processing, there is the question of how infants come to process first-order relations (i.e., categorical spatial relations—the eyes above the nose) and second-order relations (i.e., metric spatial relations—the distance between the eyes and the nose). The expertise that adults have for processing faces is believed to be associated with sensitivity to second-order relations. Initial sensitivity to both first- and second-order relations is present in infancy, although full development of sensitivity to second-order relations to adultlike levels may follow a protracted course lasting even into adolescence.

There is also a growing literature on how infants come to process social attributes of faces during their first year (e.g., identity, emotion, gender, race, attractiveness). This literature initially focused on the question of mother–stranger differentiation with its implications for the development of attachment. For example, past studies have examined how soon after birth infants display a preference for mother over stranger, how much exposure to the mother's face is needed to elicit a preference, what perceptual cues mediate the preference, and whether the preference can be manifest in other modalities besides vision. This research collectively suggests that the preference for mother over stranger is manifest in the late third trimester in the auditory domain and shortly after birth in the visual and olfactory domains. The visual preference is facilitated by increased exposure to the mother's face and voice in the first few hours and days after birth. The preference also relies increasingly on the internal features of the mother's face during the first months of life.

A related line of inquiry has investigated issues related to how infants process emotion information from faces. Studies have examined the roles of static versus dynamic cues in the extraction of emotion information from faces and how multimodal information from face and voice may contribute to infants' developing understanding of emotion. Other research has considered whether infants (1) extract emotion information more robustly when familiar individuals present that information, (2) recognize emotional expressions as members of a common category (e.g., happiness), and (3) display spontaneous looking preferences for some facial depictions of emotions over others. Taken together, the studies suggest that infants may process emotion information more efficiently from dynamic and familiar faces, and that multimodal information may contribute to infants' developing understanding of the "meaning" of emotion by lessening the likelihood that attention will be focused on modality-specific cues (e.g., toothiness in the visual input). In addition, infants can categorize emotional expressions across variation in the intensity of the emotion and the individuals depicting the emotion, and also display differential responsiveness to classes of emotion through spontaneous preference (e.g., fearful faces are preferred to happy ones).

Recent investigations have focused on how infants respond to gender and race information in faces. These studies have produced evidence consistent with the observation of familiarity preference in the face identification literature. In particular, by three months of age, infants prefer the gender of the primary caregiver and same- to other-race faces, with both preferences driven by differential experience. Experience also affects infants' face recognition memory. Specifically, three-month-olds reared by a female caregiver and presented with a series of female faces preferred a novel over familiar female face; however, when presented with male faces, there was no differential preference for a novel over familiar male face. In addition, although three-month-old Caucasian infants exposed predominantly to Caucasian faces performed as well on a recognition memory task involving either own- or other-race faces, nine-month-old Caucasian infants demonstrated recognition memory only for Caucasian faces. The recognition advantage for same-race faces and its time course of development has also been observed for human infants viewing same- versus other-species of faces (humans versus monkeys). All of the results suggest that experience with faces in the first half-year of life narrows infants' face representation from a general to a specific one that is tuned to the attributes of frequently encountered face categories.

Another social dimension of faces that infants respond to is physical attractiveness. In particular, infants will spend more time looking at attractive faces (as judged by adults) when these are shown paired with less attractive human faces. Infant preference for attractive faces has been observed for a range of human faces, including Caucasian and African American adult female faces, adult male faces, and infant faces. The attractiveness effect can be demonstrated even in newborn infants: It is orientation dependent, occurring for upright but not inverted faces, and it is driven by the internal features of faces.

A question of interest is whether the attractiveness preference in infants is dependent on perceptual learning mechanisms or whether it reflects the face representation that newborn infants bring to the learning situation for faces. The learning account of the attractiveness effect is couched in terms of an averaging process known as prototype formation: When several faces are averaged, adults perceive the resulting face as more attractive than any of the individual faces. By this learning account, infant preference for attractive faces may reflect a preference for faces similar to a composite of the faces seen since birth. This account can apply even to the results obtained with the newborn infants, given that those infants were two to three days old at testing, and would likely have experienced a multitude of faces even during that short time frame. In contrast, by a nativist account, newborn infants could enter the world with a face representation, and attractive faces are preferred because they more closely match this representation. This representation could still be in the form of a prototype, except that it would have been formed through evolutionary mechanisms. Consistent with the nativist account, young infants have also been found to prefer attractive over unattractive nonhuman animal faces for which they had no previous experience (i.e., cats, tigers).

The finding that the attractiveness preference in infants extends beyond conspecifics also suggests that it is not reflective of an adaptation to mate choice, as some have suggested, but may point toward the operation of more general mechanisms that process a family of preferred perceptual features that includes, but may not be limited to, particular features, such as large eyes, and the complex geometric attributes that characterize the spatial relations among the features, such as their location (e.g., height) and arrangement (e.g., symmetry) within the whole. Thus, just as the perception of some social attributes of faces (i.e., identity, emotion, gender, and race) seems to be driven by experience, the perception of other social attributes of faces (i.e., attractiveness) may be determined by the initial settings of our perceptual systems.

Development of Face Perception Beyond Infancy

Beyond infancy, the development of face perception seems both constrained by experiences occurring during infancy for the processing of some kinds of information and characterized by flexibility with regard to the processing of other kinds of information. For example, the absence of visual input in early infancy as a result of congenital cataracts adversely affects sensitivity to second-order, but not first-order relational information. Also, with respect to species, the ability to individuate human faces is maintained in later development, whereas the ability to individuate monkey faces, while comparable to individuating human faces before nine months of age, is now absent in adulthood. In addition, children who have suffered a history of physical abuse perceive anger in emotionally ambiguous faces more often than children who have not been abused. By contrast, although face processing seems to tune into gender and race during infancy, there is flexibility thereafter. In the case of gender, same-sex preferences and recognition advantages emerge in childhood, and there is dissociation of preference and recognition in adulthood (i.e., preference for opposite-sex faces and superior recognition for same-sex faces). Similarly, for race, Korean adults adopted by French families during childhood (ages three–nine years) display a recognition deficit for Korean faces relative to their ability to recognize European faces. Evidence such as this indicates that experience contributes to the development of face perception beyond the first year of life.

Paul C. Quinn, Kang Lee,
Olivier Pascalis, and Alan M. Slater

See also Attractiveness; Experience-Dependent Plasticity; Face Perception; Face Perception: Physiological; Infant Perception; Infant Perception: Methods of Testing; Nature and Nurture in Perception; Perceptual Development: Imitation; Perceptual Development: Intermodal Perception; Social Perception

Further Readings

Bhatt, R. S., Bertin, E., Hayden, A., & Reed, A. (2005). Face processing in infancy: Developmental changes in the use of different kinds of relational information. *Child Development, 76,* 169–181.

Cassia V. M., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science, 15,* 379–383.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science, 18,* 1084–1089.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2003). Early visual experience is necessary for some—but not all—aspects of face processing. In O. Pascalis & A. Slater (Eds.), *The development of face processing in infancy and early childhood: Current perspectives* (pp. 99–117). Huntington, NY: Nova Science.

Montague, D. P. F., & Walker-Andrews, A. S. (2002). Mothers, fathers, and infants: The role of person familiarity and parental involvement in infants' perception of emotion expressions. *Child Development, 73,* 1339–1352.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review, 98,* 164–181.

Pollak, S. D., & Kistler, D. J. (2002). Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proceedings of the National Academy of Sciences USA, 99,* 9072–9076.

Quinn, P. C., Kelly, D. J., Lee, K., Pascalis, O., & Slater, A. M. (2008). Preference for attractive faces in human infants extends beyond conspecifics. *Developmental Science, 11,* 76–83.

Robbins, R., & McKone, E. (2007). No face-like processing for objects of expertise in three behavioural tasks. *Cognition, 103,* 34–79.

Sai, F. Z. (2004). The role of the mother's voice in developing mother's face preference: Evidence for intermodal perception at birth. *Infant and Child Development, 14,* 29–50.

Sangrigoli, S., Pallier, C., Argenti, A. M., Ventureyra, V. A. G., & de Schonen, S. (2004). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science, 16,* 440–444.

PERCEPTUAL DEVELOPMENT: HEARING

Infants respond to sounds of many types from birth. In fact, *hearing* begins months before a child is born. However, many studies demonstrate that hearing continues to develop well into adolescence. The development of hearing, as will be shown in this entry, depends on improvements in the way that sound is initially represented in the nervous system, in the way that the auditory scene is organized, in the ability to select sounds for processing, and in the flexibility of sound processing.

Sound

All sounds can be described in terms of frequency, level, and changes in frequency or level over time. Frequency is the acoustic dimension that gives rise to the percept of pitch. Level is the acoustic dimension that gives rise to the percept of loudness. Changes in frequency and level over time help to identify sound sources and to determine which sounds are being produced by the same source, among other things. The ear represents a sound's spectrum, the level of sound at each frequency, as well as its waveform, the sound amplitude over time. This representation must be faithfully conveyed through the auditory nervous system for sound to be accurately perceived.

Development of the Ear

The inner ear, specifically the cochlea, is responsible for converting sound into a neural response that represents the spectrum and waveform of sound. The cochlea begins to respond to sound—for the most part sounds generated by the mother—late in the second trimester of gestation. By the last month of gestation, every indication is that the cochlea functions much as it does in adults. Numerous studies have shown that newborn infants recognize their mother's voice and other things about the sounds to which they were exposed prenatally.

For sound to get into the inner ear postnatally, however, it must be conducted through the external and middle ears. The external ear includes the pinna and the ear canal. The external ear not only leads sound into the head but also shapes the spectrum of sound. The result is that frequencies in a particular range—between 2000 and 5000 hertz (Hz) in adult males—will be delivered to the middle ear at higher levels than are other frequencies. Infants' smaller pinnas and shorter ear canals will tend to resonate at higher frequencies than the adult ear.

The middle ear contains the eardrum and the ossicles, the small bones that connect the eardrum to the inner ear. It makes it possible for airborne sound to stimulate the structures in the fluid-filled inner ear. The efficiency with which sound is conducted through the middle ear is reduced in infants. In newborns, the combined effect of the immature external and middle ears is a small reduction in sound conduction of frequencies below 2000 Hz and a larger reduction in sound conduction in frequencies above 2000 Hz. High-frequency sound conduction improves dramatically during infancy, but conductive efficiency is not completely mature until 10 to 11 years of age.

Auditory Neural Development

Immature transmission of information through the auditory brain stem also limits hearing during infancy. It appears that auditory neural transmission is inefficient early in infancy, resulting in less information about sound reaching higher levels of the brain. Information is also transmitted more slowly and less synchronously in the immature auditory brain stem. Neural responses at higher levels continue to develop well into childhood, although the role that these play in hearing development is not clear.

Auditory Behavior

Early Infancy

This conductive and neural immaturity is reflected in infants' and children's auditory behavior. In infants younger than six months of age, it appears that the representations of a sound's spectrum and waveform are both immature. Although even newborns respond to sound changes along any dimension, they require a bigger change in a sound before they respond. A three-month-old infant, for example, has poorer resolution of high frequencies, resulting in particular difficulty in discriminating between different high-frequency tones. Infants have higher hearing thresholds than adults across the range of hearing, but during the first six months of life, infants' thresholds are particularly poor for frequencies above 2000 Hz. In these early months, infants only hear rather large changes in a sound's level, and they have greater difficulty hearing one sound when it occurs very shortly after another sound. Even young infants have sufficient sound resolution to allow them to discriminate between speech sounds and to recognize voices, musical instruments, and simple melodies, albeit with less precision than an adult.

Later Infancy and Preschool Period

As a result of improvements in conductive and neural efficiency, hearing improves in several ways by six months of age. A six-month-old's ability to resolve high frequencies is much like an adult's. Sounds that occur close together in time can be resolved. Hearing thresholds are still elevated by 10 to 15 decibels (dB) relative to adult thresholds across the range of hearing, however, and though they improved over that at three months, the ability to respond to a change in a sound's level remains immature.

Mature frequency and temporal resolution allows the infant to process sound in a more sophisticated fashion. By the middle of the first year of life, infants are beginning to become attuned to some of the specific acoustic characteristics of native speech. Over the next six months, the ability to process native speech sounds improves, while the ability to process some nonnative speech sounds declines. For example, infants learning to speak Japanese lose the ability to distinguish English "r" from "l," and infants learning to speak English lose the ability to distinguish Hindi retroflex and dental stops. Seven-month-olds recognize the pitch of a complex sound despite changes in timbre; for example, an infant can recognize that a piano tone and a violin tone have the same pitch, even though the tones have a different quality.

Infants and preschoolers localize sounds in space with much poorer precision than adults. The development of sound localization depends on the acoustics of the head and external ear, the neural circuits that compare the inputs between ears, and the formation of a "map" that relates acoustic cues to locations in space. The external ear's shaping of the spectrum of sound varies systematically with a sound's location in space. A sound's elevation can be determined on this basis. The most important cues for localizing a sound in the horizontal plane are differences in the way that the sound's spectrum is shaped at the two ears and differences in the arrival time of the sound at the two ears. These are referred to as interaural cues. A small external ear and head shape sound differently and provide small interaural cues. Moreover, the interaural cues specifying most locations in space will continually change as the head and ears grow. However, neither limited acoustic cues nor immature neural processing of interaural cues can fully account for infants' poor sound localization. It is possible that the maturation of sound localization also depends on the ability to map acoustic cues onto positions in space. The effects of interaural cue disruption on later sound localization are consistent with that hypothesis. The accuracy of sound localization approaches adult values around five years of age, although older children have difficulty localizing sounds when other sounds or reverberation are present.

In contrast to an infant, a four- or five-year-old's ability to hear level changes is mature and hearing thresholds are within five dB of adult values, limited only by conductive efficiency. By this age, the relationship between a sound's level and its loudness is like that seen in adults. According to some research, the improvement in both hearing threshold and in the representation of level during later infancy and the preschool period results from a reduction in response variability in the nervous system. Other research indicates that hearing matures as a result of improvements in the ability to separate sounds and the ability to process one sound among other sounds. Immaturity of these abilities is evident if one compares infants and children to adults in tone detection. If tones of very different frequencies are presented simultaneously, an adult will be able to detect any of the individual tones at as low a level as when that tone is presented alone. If an adult is listening for a near-threshold tone at one frequency, near-threshold tones at very different frequencies will not be heard. Thus, the adult is "listening" at a specific target frequency in these tasks. Infants and young children, in contrast, will have higher thresholds for a tone if other tones are present, despite the fact that they are able to resolve the tones. Furthermore, if an infant or child expects that a near-threshold tone at one frequency will be presented, near-threshold tones at very different frequencies will still be heard. Thus, infants and children do not listen at a specific target frequency. Because there is always noise present—in the environment, ear, or neural response—failure to listen in a selective way will influence thresholds whether or not other sounds are intentionally presented.

Limited ability to separate a complex sound into different sound sources, known as auditory scene analysis, may contribute to infants' and children's tendency to listen broadly. In real sound environments, this is the ability that allows a listener to organize the auditory scene: a bus straight ahead, a man speaking to the right, music from a car passing by on the left. Both acoustic and visual cues provide the basis of auditory scene analysis. Under conditions in which adults can separate sounds with minimal cues, providing additional cues improves infants' and children's ability to separate one sound from another. This suggests that immature auditory scene analysis is related to immature listening styles.

Infants' and preschoolers' "broad" listening style also influences their speech perception. There are multiple acoustic cues that distinguish one speech sound from another. Although adults tend to weigh some cues more heavily than others in identifying speech sounds, manipulation of any cue tends to change adults' identification of speech predictably. Preschool children, on the other hand, use a more broad-brush approach, identifying speech sounds on the basis of global properties rather than fine details. This may result in errors in children's speech identification when the sounds are very similar or when listening in a noisy environment.

School-Age Children

A six- or seven-year-old has achieved a detailed representation of sound and is able to use the details in that representation to identify sounds. Although in many respects, older children process sound as adults do, complex listening situations may still make it difficult for children to do so. Situations in which several people are speaking simultaneously, particularly if the voices or speech content are similar, create a greater listening problem for children than for adults. Under difficult listening situations, the acoustic cues that allow a listener to identify speech may differ from the cues that are used in easy listening conditions. Adults are able to weigh the acoustic cues differentially under different listening conditions. It has been suggested that although children have access to the details of complex sound, they do not yet have the flexibility to switch from one cue to another as listening conditions change. Even adolescents do not seem to be as facile as adults in this regard.

Lynne A. Werner

See also Audition; Auditory Frequency Selectivity; Auditory Localization: Psychophysics; Auditory Masking: Auditory Scene Analysis: Auditory Thresholds; Speech Perception

Further Readings

Buss, E., Hall, J. W., & Grose, J. H. (2006). Development and the role of internal noise in detection and discrimination thresholds with narrow band stimuli. *Journal of the Acoustical Society of America, 120,* 2777–2788.

Keefe, D. H., Burns, E. M., Bulen, J. C., & Campbell, S. L. (1994). Pressure transfer function from the diffuse field to the human infant ear canal. *Journal of the Acoustical Society of America, 95,* 355–371.

Litovsky, R. Y., & Ashmead, D. H. (1997). Development of binaural and spatial hearing in infants and children. In R. H. Gilkey & T. R. Anderson (Eds.), *Binaural and spatial hearing in real and virtual environments* (pp. 571–592). Manwah, NJ: Lawrence Erlbaum.

Moore, J. K., & Linthicum, F. H. (2007). The human auditory system: A timeline of development. *International Journal of Audiology, 46,* 460–478.

Schneider, B. A., Trehub, S. E., Morrongiello, B. A., & Thorpe, L. A. (1989). Developmental changes in masked thresholds. *Journal of the Acoustical Society of America, 86,* 1733–1742.

PERCEPTUAL DEVELOPMENT: IMITATION

Imitation is a powerful learning mechanism in human beings, and it is a capacity that begins early in childhood. Before children learn through verbal instructions, they learn through imitation. A three-year-old will rummage through her mother's purse to find lipstick to apply to her face. Children crawl up to their parents' computers and poke the keys. They do these things despite being told that they should not, suggesting that imitation is not due to Skinnerian conditioning. Nor are these behaviors the result of independent invention and chance motor movement. These everyday events illustrate a basic human capacity: Children perceive others' actions and are motivated to imitate what they see. Other animals imitate in rudimentary ways, but scientists agree that *Homo sapiens* is the most imitative species in the animal kingdom. This entry defines and gives the functions of imitation, and relates imitation to the areas of childhood, neuroscience, autism, and social robotics.

Definition

Imitation occurs when three conditions are met. It requires that the imitator: (1) perceives the act of another, (2) repeats that act, and (3) does so with the goal of matching what was perceived. Accidental mirroring (e.g., people slipping on the same icy street) is not imitation. Imitation requires perception of the model; and it requires that the perceived act becomes a target or goal the imitator is striving to match.

Childhood Imitation

Traditional theories of human development underestimated human infants' capacity to imitate. In the classic developmental theory of Jean Piaget, there were six stages of imitative development. Piaget postulated that young infants lacked the cognitive capacity to match facial gestures. More recent experiments established that newborns can imitate simple facial gestures, such as poking out their tongues and opening and closing their mouths. Detailed motor analyses show that they correct their behavior, honing in on an accurate match. This indicates that infants are using cross-modal perception to imitate. The adult body movement is the target, and infants use proprioceptive feedback from their own unseen movements to guide their motor system to match the movement of the visual target.

Imitation would be of limited utility if it was restricted to direct mimicry of events in the immediate perceptual field. Research shows that infants can imitate from memory. In one study, 14-month-old infants saw an adult perform a novel action, but were not given the object. They returned to the laboratory 24 hours later and were handed the toys. Videotaped records showed imitation of the absent act from memory. Children not only imitate real people, but images on television. However, infants are poorer at imitating from television. It may be that the television is not sufficiently social to elicit maximal imitation; alternatively children have a difficulty using a two-dimensional (2-D) representation to guide their three-dimensional (3-D) actions, a transfer of learning problem.

After infancy, television becomes a potent source of models for children, and there is concern about learning violence from perceptual exposure via television. Older children learn not only concrete motor behaviors, but also more abstract information from television and live models. For example, preschool children preferentially imitate same-sex models. Thus, imitation is a mechanism for learning gender roles—how girls (or boys) act in a particular culture.

Functions of Imitation

Imitation has three functions. First, it serves a social function. In infancy, imitation games establish reciprocity and communication between caretaker and child. Even in adults, action mirroring is bound with communication. Psychotherapists "mirror" their patients' actions or postures conveying empathy and caring. Thus, perceptual-motor coupling has social-emotional value: People are comforted by behavioral resonance with others.

Second, imitation is a vehicle for the spread of culture. Just as DNA transfers biological information, imitation is a mechanism for transferring newly invented actions, characteristics, and tool use across generations. Without imitation, each generation of early humans would have had to "reinvent the wheel" rather than profiting from watching the practices of older generations. The cultural accumulation of knowledge, traditions, and social roles is called "cultural learning" or "cultural evolution" (to distinguish it from Darwinian evolution), for which there is no counterpart in other species.

Third, imitation may help the child understand another person's motivations and internal states. Human children develop the understanding that others have mental states—such as beliefs, desires, and emotions—that differ from their own. This is called a "theory of mind." The mental states of others cannot be seen, touched, smelled, tasted, or heard. Understanding others' minds is not innate. How is it learned? Imitative processes provide a developmental mechanism. We first learn the conjunction between our own mental states and behavior. When we feel happy (internal state), we smile (behavior). When we intend to pick up an object, we reach out our hand. The ability to imitate provides the opportunity for children to compare their own actions to similar actions of others. This is serves as a foundation for the realization that others are "like me" in deeper ways as well: Others who *act* like me may also *feel* like me.

Neuroscience, Autism, and Social Robotics

Scientists are beginning to explore the neural substrates that support imitation. Current data suggest that the inferior parietal lobe and the medial prefrontal cortex are activated in adult imitation tasks. The involvement of the medial prefrontal cortex in imitation is noteworthy because it is also activated in high-level theory-of-mind tasks. Some scientists refer to the neural networks underlying human imitation as the "shared neural systems for perception and action" or the "human mirror neuron system."

Children with autism have profound deficits in imitation. Scientists are now seeking to use motor imitation deficits as an early marker in the diagnosis of autism. Studies suggest that intensive therapy on social imitation games may boost not only simple mimicry but also a much broader set of skills in social understanding.

The robotics community is trying to design robots that are able to "learn by example," like human children. In the future, if a robot's sensors could perceive a human action and translate it into the robot's "motor" system, the robot might be able to watch tutors and learn to pour a cup of tea, tie a shoelace, or lift a heavy load onto a stack. Such open-ended "learning by watching" routines are beyond current robots, but many philosophers and computer scientists believe that the first step toward a social robot will be to crack the imitation problem and have the robot learn like a baby—through imitation.

Andrew N. Meltzoff

See also Perceptual Development: Intermodal Perception; Social Perception; Theory of Mind

Further Readings

Hurley, S., & Chater, N. (Eds.). (2005). *Perspectives on imitation: From neuroscience to social science* (Vols. 1 & 2). Cambridge: MIT Press.

Meltzoff, A. N., & Prinz, W. (Eds.). (2002). *The imitative mind: Development, evolution and brain bases.* Cambridge, UK: Cambridge University Press.

Nehaniv, C. L., & Dautenhahn, K. (Eds.). (2007). *Imitation and social learning in robots, humans, and animals: Behavioural, social and communicative dimensions.* Cambridge, UK: Cambridge University Press.

Piaget, J. (1962). *Play, dreams and imitation in childhood.* New York: Norton.

PERCEPTUAL DEVELOPMENT: INFANT MUSIC PERCEPTION

The world of infants is filled with music. Around the world caregivers sing to infants, rousing them with play songs and lulling them to sleep with lullabies. In Western culture, toys and programming aimed at infants often include music. Even when people talk to preverbal infants they add musical features to their speech, such as higher pitch, wider pitch contours, highly rhythmic patterning, and abundant repetition. Some questions discussed in this entry are: Why is music for infants so prevalent? What do infants perceive when they hear music? How does musical acquisition compare to linguistic acquisition? And how does musical development interact with other aspects of perceptual and cognitive development?

As with languages, there are many different musical systems, such as Western tonal structure and Indian raga structure. As with languages, there appear to be a number of near universal features of musical systems. These include hierarchical metrical organization (the perceived underlying beat and its sub- and superdivisions), the functional equivalence of tones an octave apart (doubling of fundamental frequency or pitch), the use of scales consisting of a small number (usually between five and nine) of notes per octave (although the particular intervals between scale notes differ across systems), and the central role of the consonance/dissonance dimension in the ebb and flow of musical tension and relaxation. Musical systems also differ from each other in features such as the complexity of the metrical structure, the particular scales that are used to form melodies, and whether or not harmony or a drone or a monophonic texture is used. Infants acquire culture-specific implicit knowledge about the music system(s) to which they are exposed, just as they acquire the language(s) to which they are exposed.

Infant Preferences

Infants appear to enjoy listening to music. Unlike monkeys, who prefer silence to music, infants prefer music to language and prefer music to silence. The origin of this affinity for music remains unknown. However, some evolutionary psychologists have argued that music evolved in order to enhance the emotional bond between infants and their caregivers. Evidence consistent with a central role for emotion in infants' perception of music comes from studies of infant preferences. Typically these are tested in a preferential looking paradigm. A light on one side of the infant flashes until the infant looks to that side, at which time the light stays on, illuminating a toy, and one sound stimulus begins to play. As long as the infant looks at the toy, the sound stimulus continues to play. Once the infant looks away, the light extinguishes, the toy disappears and the sound stimulus stops. The second trial begins with a light flashing on the other side. When the infant looks to this side, again the light stays on, illuminating the toy, and a second sound stimulus begins to play. Again it plays until the infant looks away. A number of both types of trials are presented, with the infant controlling how long he or she listens to each trial. If infants listen significantly longer to one stimulus over another, it can be inferred that they prefer this stimulus. Studies of infant preferences reveal that they prefer to listen to infant-directed renditions of a song compared to adult-directed renditions. Across cultures, infant-directed renditions are typically higher in pitch, slower in tempo, and rendered in a more loving tone of voice. The latter quality appears to be particularly important to infants—the degree to which they prefer the infant-directed over the non-infant-directed version is correlated with the degree to which the infant-directed version is rated as more loving in tone by adults.

Caregivers also appear to use different types of infant-directed singing to achieve different caretaking goals. Even in Western culture where CDs are readily available, diary studies indicate that parents sing to infants in many contexts, such as during play, during diaper changes, while eating, while driving in the car, and while going to sleep. Two basic styles of singing have been identified: lullabies and play songs. These song types are readily discriminated and categorized by adults, and can also be distinguished on the basis of acoustic features. Infants react differently to the two types of infant-directed singing, for example, focusing their attention outward during play songs, and looking downward during lullabies, suggesting that the different styles of singing serve different functions related to maintenance of infant state. In sum, singing to infants is a universal behavior, and infants respond particularly to the emotional aspects of infant-directed singing.

Enculturation to Pitch Structure

In general, young infants are sensitive to universal aspects of musical pitch structure, which likely involve innate constraints on auditory processing in the nervous system. One near-universal feature of musical systems is the importance of consonant intervals. Two tones that sound pleasant or smooth together are said to be consonant, whereas tones that sound unpleasant or rough together are said to be dissonant. Complex tones with pitch typically have energy at a fundamental frequency (that of the perceived pitch) and at harmonics at integer multiples of the fundamental. The fundamental frequencies of the tones forming a consonant interval stand in small-integer ratios (e.g., 1:2, 2:3), whereas those of dissonant intervals stand in more complex ratios (e.g., 9:11; 32:45). The continuum between consonance and dissonance is perceived similarly around the world, although the point at which the category is perceived to change from consonance to dissonance varies across cultures. Consonant and dissonant intervals are already processed differently in the auditory periphery. The basilar membrane in the inner ear encodes frequency content along its length. Dissonant intervals contain harmonics that are sufficiently close in frequency as to interact on the basilar membrane. They also set up distinct neural firing patterns in the auditory nerve. Young infants (two months old) have been tested, and they can discriminate and categorize consonant and dissonant intervals. Furthermore, like adults, infants prefer to listen to consonant over dissonant tone pairs, and prefer to listen to a Mozart sonata in the original compared to a version with occasional semitone changes that introduce additional dissonance. Interestingly, although many species discriminate consonant and dissonant intervals, only humans appear to prefer consonance over dissonance.

Most musical scales contain prominent consonant intervals, and tones an octave apart (e.g., two C's or two D's in Western musical notation) are considered functionally equivalent. At the same time, musical scales differ in the particular intervals they contain. Through everyday exposure to a musical system, infants gradually acquire musical system-specific knowledge, just as they acquire a specific language. Western adults, even those without musical training, readily detect wrong notes in a melody if the notes go outside the scale or key on which the melody is based. However, they are much poorer at detecting wrong notes that remain within the key, especially when they do not change the implied harmony (the chords that would be used to harmonize the melody). Using a conditioned head turn method in which infants are rewarded with dancing toys for turning their head when a wrong note is played, it has been shown that eight-month-olds can detect changes that maintain the key and harmonic structure just as well as they detect changes that violate the key structure and harmonic structure. Infants can actually outperform adults in some cases on changes that remain within the key. Though this seems impressive, what it really indicates is that infants have not yet learned to process melodies according to Western scale or harmonic structure. Harmonic processing appears to have a particularly protracted development, with sensitivity emerging around four to six years of age in the absence of formal musical training, and with adult levels not achieved until several years later.

Enculturation to Rhythmic Structure

Rhythm is arguably the most fundamental aspect of music: It is possible to compose music without melody, but it is difficult to do so without temporal structure. From at least as young as two months, infants can discriminate different rhythmic patterns, and they show optimal tempo discrimination at the same speed as adults (100 beats per minute). Musical rhythm and movement are intimately connected. When listening to music, people perceptually extract an underlying regular hierarchical beat structure (whether or not there are actual sound events on every beat), called the metrical structure. The meter indicates the points at which to tap to the beat, as well as how to dance to the music. Not only does music make us want to move, but how we move also affects the metrical structure that we perceive, even in infants. Specifically, when infants are presented with an ambiguous repeating rhythm (i.e., with no physical accents) as they are bounced up and down on every second beat (as in a march), they subsequently prefer to listen to a version of the rhythm with accents every second beat compared to a version with accents every third beat (as in a waltz). However, if they listen to the identical rhythm as they are being bounced on every third beat, they subsequently prefer the version with accents every third beat. Movement involves a number of systems, such as motor planning, tactile, proprioception, and vestibular systems. Interestingly, the vestibular system, which gives us our sense of balance and informs us about how we are moving relative to the gravitational field, appears to be important for the relation between auditory meter and rhythmic movement. The connection between movement and auditory rhythm emerges very early in development, and rhythmic movement evolved much earlier than auditory rhythm perception (e.g., even a jellyfish moves rhythmically), suggesting that the perception of auditory rhythm may have evolved from circuits for rhythmic movement.

There is considerable diversity across cultures in the complexity of rhythms used in music. Western meters tend to be fairly simple, with time intervals typically divided into groups of two or three beats. However, music from many cultures involves considerably more complexity. Enculturation to rhythm develops in infancy. Adults from Bulgaria and Macedonia have no trouble perceiving the complex meters that occur in the folk music of their cultures, but Western adults have great difficulty detecting disruptions to these complex meters. Interestingly, Western infants of 6 months of age also have no difficulty detecting disruptions to both simple and complex meters, but by 12 months of age, Western infants detect only disruptions to simple meters.

Formal Musical Training

Young infants are initially sensitive to universal aspects of musical pitch and rhythmic structure. Through exposure to a particular musical system, infants internalize the statistical regularities in the input, and their perceptual system becomes specialized for processing the particular pitch and rhythm structures of the musical system of exposure. In older children, formal musical lessons, as opposed to simple passive exposure to music, enhance brain development in auditory areas, and the development of attention, memory, and general IQ. It currently remains unknown whether formal musical training in infancy would confer similar advantages. However, there are enormous changes in the auditory cortex during the first months after birth, suggesting that the neural circuits for processing music might be particularly plastic during this time.

Laurel J. Trainor

See also Language; Music Cognition and Perception; Perceptual Development: Hearing; Perceptual Development: Speech Perception

Further Readings

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences, 11,* 466–472.

Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science, 16,* 48–55.

Patel, A. D. (2008). *Music, language and the brain.* Oxford, UK: Oxford University Press.

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science, 308,* 1430.

Trainor, L. J. (2005). Are there critical periods for music development? *Developmental Psychobiology, 46,* 262–278.

Trainor, L. J. (2008). The neural roots of music. *Nature, 453,* 598–599.

Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 394–402.

Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience, 6,* 669–673.

PERCEPTUAL DEVELOPMENT: INTERMODAL PERCEPTION

Speaking faces, baking bread, speeding cars—the world provides a richly structured, continuously changing stream of stimulation to all of our senses. *Intermodal perception* (also called intersensory or multimodal perception) refers to perception of information from objects or events available to multiple senses simultaneously. Because most objects and events can be seen, heard, and touched, everyday perception is primarily intermodal. Despite the fact that information about the world is carried through different sensory channels that each provide distinct forms of stimulation, we are able to perceive a stable world of unitary objects and events (people speaking, cars honking), rather than separate sights, sounds, and tactile impressions. The senses work together as a coordinated perceptual system, even in newborns, and intermodal perception develops rapidly and with increasing specificity across infancy. How we accomplish this "integration" is a puzzle that has fascinated philosophers and scientists for centuries. This entry covers the history and theory of intermodal perception, as well as aspects of the development of the senses as they relate to intermodal perception.

History and Theory

Over 2,000 years ago, Aristotle proposed a *sensus communis,* or common sense, to explain how we perceive qualities common to different senses, such as number, form, and unity. Centuries later, philosophers such as John Locke and George Berkeley proposed that we needed to integrate information across separate sensory channels before perceiving a unified object or event, such as a bell ringing, rather than perceiving separate streams of light and sound. This posed a "binding" problem for perception, in that the brain would need to somehow unify separate channels of sensory information to "construct" a coherent world. This constructivist view dominated thinking throughout most of the 20th century, including Jean Piaget's well-known theory of cognitive development. Piaget proposed that integration took place gradually over the first year of life through interacting with objects and coordinating information across the senses. Prior to integration, infants were thought to perceive a world of unrelated sights, sounds, smells, and tactile impressions, much like the "blooming, buzzing confusion" described by William James in 1890.

Not until the psychologist James J. Gibson published his work on ecological perception in 1966, *The Senses Considered as Perceptual Systems,* was this constructivist view seriously questioned. Gibson argued that our senses work together as a unified perceptual system, and that the existence of different forms of sensory stimulation was an advantage rather than a disadvantage for perception of unified objects and events. This is because we perceive amodal (not tied to a particular sense modality) information that is redundant or identical across the senses. This includes temporal synchrony, rhythm, duration, tempo, changes in intensity, and spatial location common to audiovisual events, and shape, substance, size, and texture common to visual–tactile events. Because all events occur over time and space, all events provide amodal information.

Gibson's theory is considered a differentiation view of development. In contrast to the integration view, which holds that the senses are separate at birth, the differentiation view, articulated by Eleanor J. Gibson, proposes that the senses are unified at birth and that perceptual development is characterized as a gradual process of differentiating increasingly finer levels of sensory stimulation. From this view, now widely accepted among developmental psychologists, there is no "blooming, buzzing confusion" in early development. Rather, infants detect many general features of unified multimodal events from birth and learn to perceive increasingly more subtle differences and more complex objects and events through looking, listening, and interacting with their environment.

Importance of Amodal Information and Intersensory Redundancy for Development

How do infants learn to detect unified multimodal events such as a person speaking? Evidence indicates that amodal information, particularly temporal synchrony between sights and sounds, provides the glue that binds information across the senses and thus serves as a solution to the age-old binding problem. For example, by detecting the synchrony, rhythm, and tempo common to the sights, sounds, and tactile impressions of one's hands clapping, we attend to the event as a whole, rather than separate sights, sounds, and tactile stimulation. When the same amodal information is concurrently available and synchronized across more than one sense modality, this is called *intersensory redundancy*. Research has shown that intersensory redundancy is highly salient to humans and animals. For example, the special attention shown to social stimulation by infants of many species is likely a result of the high degree of intersensory redundancy social partners provide. Further, we know that redundancy highlights unitary multimodal events and their amodal properties, at the expense of other properties, much like a figure-ground effect. For example, an infant will notice the intersensory redundancy between the face and voice of a person speaking (synchrony, rate, rhythm, and co-location of audiovisual speech) at the expense of other (modality-specific) information, such as the appearance of the face or the pitch and timbre of the voice. In contrast, when redundancy is not available, such as when a person is still and silent, infants preferentially explore nonredundant (modality-specific) aspects, such as the appearance of the face.

Thus, redundant, amodal information plays a significant role in guiding and constraining which aspects of events we selectively attend, particularly in early development when attention is least flexible and capacity is most limited. Later, when attention is more flexible and efficient, we can attend to amodal properties as well as more specific properties of the same events in a shorter time. Furthermore, because selective attention provides the basis for what is perceived, learned, and later remembered, intersensory redundancy has a powerful organizing influence on early perceptual, cognitive, social, and emotional development.

Development of Auditory-Visual Perception

Much research has focused on the ability of young infants to perceive multimodal events, and in particular, information that is amodal and redundant across audition and vision. Scientists have discovered that even young infants are skilled at perceiving amodal information, that intermodal perception improves across the first year of life, and that it develops in order of increasing specificity, with global information detected developmentally prior to more specific levels of stimulation, consistent with the pattern of differentiation and "increasing specificity" proposed by Gibson.

Audiovisual Space

As early as the first weeks of life, infants reliably move their eyes in the direction of a sound. This early coordination of auditory and visual space is important because it enables infants to discover visual information at the source of the sound and thus promotes detection of intersensory redundancy. At first, localization is rather imprecise, but it improves rapidly across infancy. Through auditory and visual experience with objects and events, infants calibrate and align their sense of audiovisual space.

Object and Event Perception

In the first months following birth, infants can detect the temporal synchrony and spatial co-location uniting the sights and sounds of an object moving and striking a surface. By two to five months, infants match soundtracks with object motions on the basis of finer levels of amodal information, such as substance (elastic vs. rigid), composition, tempo, and rhythm. Soon thereafter, infants detect arbitrary, modality-specific information, such as the relation between the color and shape of an object and the pitch of its sound. Detection of global, sight-sound relations (e.g., synchrony) promotes detection of more specific amodal relations (e.g., tempo or rhythm), and this in turn promotes detection of arbitrary, modality-specific relations (e.g., pitch and color). These remarkable abilities illustrate the principle of increasing specificity and how processing of global information such as intersensory redundancy scaffolds (guides and supports) the development of more specific processing.

Social Development

People provide a great deal of multimodal stimulation for infants, including talking, laughing, singing, and touching. Adults and infants also engage in richly structured multimodal interactions, called *protoconversation.* This highly intercoordinated mutual exchange of sounds, movements, and touch is a foundation for communication and social development. These early experiences contribute to and shape infants' fascination with social stimulation. Research has shown that newborns look longer to faces that are paired with a voice and prefer their mother's face hours after birth if they have previously seen her face together with hearing her familiar voice. By three months, infants can match facial and vocal expressions of emotion, such as happy, sad, and angry, in familiar people, and by five months, in unfamiliar people. Discrimination of emotional expressions, like perception of the rhythm and tempo of events, emerges first in multimodal (audiovisual) stimulation (by four months) and then is later extended to unimodal auditory (around five months) and unimodal visual (seven months) stimulation.

Speech Perception and Language

Speech is inherently multimodal, involving coordinated facial, vocal, and gestural information, and audiovisual redundancy promotes learning in this domain as well. For example, adults typically speak to infants using a distinctive speech pattern with slower tempo, larger pitch changes, and more repetition. This style of speaking is called *motherese,* and research has found that infants across cultures and languages are particularly attracted to this type of speech. Adults also produce exaggerated facial expressions and gestures coordinated with speech during motherese. This multimodal motherese exaggerates amodal, redundant information (intensity, rhythm, and tempo changes produced by synchronous facial and vocal stimulation), making it easier for infants to maintain attention to and learn speech. Thus, speech is first learned and experienced in the context of multimodal, redundant stimulation. Even during the last trimester of prenatal development, fetuses likely receive multimodal stimulation from the sound of the mother's voice, coordinated with synchronous body movements, vibrations of her spinal column, and movements of her diaphragm. This redundancy may contribute to the salience of the mother's voice observed soon after birth.

By the age of two months, infants can detect voice–lip synchrony during speech. By five to seven months, they can use face–voice synchrony to selectively attend to a stream of speech, separating it from a background of other sounds, and pick out individual words (speech segmentation) much better than in auditory stimulation alone. Audiovisual synchrony is also important for learning the arbitrary relation between speech sounds and the objects to which they refer, a cornerstone of language development. For example, by seven months, infants can learn to pair a speech sound with an object, provided the sound is uttered in synchrony with moving and showing the object. However, it is not until many months later that infants can learn to pair a speech sound with an object in the absence of synchrony. Parents are intuitively aware of this need and synchronize naming and showing objects to their younger infants.

Audiovisual Interactions and Illusions

Several intriguing audiovisual illusions illustrate the powerful interaction of auditory and visual perception. For example, when we view the face of a person speaking one speech sound (e.g., "ga"), while hearing a different speech sound (e.g., "ba"), we perceive another sound altogether ("da," a blend between the two). Infants show this *McGurk effect* within the first six months following birth, illustrating that infants, like adults, merge information for speech across the senses. This is also illustrated by the *ventriloquist effect.* By moving a dummy's mouth in synchrony with the ventriloquist's own speech sounds, the ventriloquist creates amodal information and fools us into perceiving that the dummy is actually speaking. In other words, although sound is coming from one location, it is "heard" as coming from another location.

Development of Visual-Tactile and Visual-Motor Perception

Amodal information also guides and constrains perception across vision and touch. For example, when we feel an object with our hands, we can perceive the same shape, size, texture, and substance that we see. One-month-old infants (and in one study, even newborns) perceive a correspondence between an object they have explored tactually (by holding or mouthing) and a visual display of the same object and prefer to look at the familiar shape. Tactile exploration and active touch are coordinated with visual exploration and develop rapidly across the first year, providing increasingly more detail about objects and their surfaces.

Proprioception is information about self-movement based on feedback from the muscles, joints, and vestibular (balance) system. This important information about the self is coordinated with motor behavior and visual feedback from motion and develops rapidly over the first year of life. By three to five months, infants can distinguish between self and other on the basis of this information. They discriminate between a live video of their own legs kicking versus those of another infant by detecting temporal synchrony and spatial correspondence between proprioceptive feedback and the visual display of their own leg motion. Infants also show *visually guided reaching,* even in the first months of life. This entails continuous adjustments in manual and reaching behavior as a function of the changing position of objects, and this intercoordination becomes more precise with experience. A little later in development, infants can adapt their crawling and early walking attempts based on visual information about the slope and solidity of the surface. Visual feedback is essential for crawling, walking, and maintaining an upright posture at all stages of development, illustrating the essential link between vision, proprioception, and motor behavior.

Odor and Vision

Young infants show recognition of their mother on the basis of her smell, and breastfed infants prefer the odor of their mother over that of another lactating woman, illustrating their sensitivity to object-odor combinations. In one study, four-month-olds looked more to an object that had previously been paired with a cherry odor when the odor was present, but not when it was absent, showing memory for the relation between the visual appearance of an object and its odor. In many species, recognition of the mother may be promoted by fetal sensitivity to the odor and taste of amniotic fluid and its similarity to other maternal odors.

Lorraine E. Bahrick and Robert Lickliter

See also Amodal Perception; Attention: Cross-Modal; Binding Problem; Direct Perception; Infant Perception; Multimodal Interactions: Visual–Auditory; Perceptual Development: Face Perception; Perceptual Development: Hearing; Perceptual Development: Object Perception; Perceptual Development: Taste and Olfaction; Perceptual Development: Touch and Pain

Further Readings

Bahrick, L. E., & Lickliter, R. (2002). Intersensory redundancy guides early perceptual and cognitive development. In R. Kail (Ed.), *Advances in child development and behavior* (Vol. 30, pp. 153–187). New York: Academic Press.

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development.* New York: Oxford University Press.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin, 126,* 281–308.

Lewkowicz, D. J., & Lickliter, R. (1994). *The development of intersensory perception: Comparative perspectives.* Hillsdale, NJ: Lawrence Erlbaum.

Rosenblum, L. D. (2005). Primacy of multimodal speech perception. In D. Pisoni & R. Remez (Eds.), *Handbook of speech perception* (pp. 51–78). Malden, MA: Blackwell.

PERCEPTUAL DEVELOPMENT: OBJECT PERCEPTION

When looking around the kitchen, most adults have no trouble determining that the toaster is separate from the blender, even if the surfaces of the two are touching. We can determine where one object ends and another begins without careful thought or effort. Is the same true for infants early in the first year of life? If so, which cues do they use to make such judgments and at what point in development are infants capable of making these judgments accurately?

Some of the first observations that provided insight into how infants perceive objects were published by Jean Piaget, who is widely regarded as one of the founders of the field of cognitive development.

He reported that his 10-month-old son Laurent would reach for a matchbox when Piaget held it out on the tips of his fingers, but as soon as it was placed on top of a large book, Laurent would reach for the book instead. These observations led Piaget to conclude that during the first year of life, infants regard object boundaries as changeable and context-dependent, rather than characteristics of objects that are lasting in time and space. This entry describes interpretations of partly occluded objects, parsing adjacent objects, and implications for infants' perceptual development as they relate to object perception.

Interpretations of Partly Occluded Objects

These observations inspired researchers to conduct more systematic research on infants' object perception. For example, in one set of studies, Phillip Kellman and Elizabeth Spelke tested how four-month-old infants perceive a display consisting of a rod with its center hidden behind a screen so that only the top and bottom of the rod are visible. Would infants interpret the visible portions of the rod as connected behind the screen, or would they expect these rod parts to be disconnected?

To answer this question, the authors measured infants' looking time as they watched the partly occluded rod display, continuing to show this same display until the infants' looking time had decreased to about half of the original level. Then, in test, the infants were shown two possible interpretations of the rod without the screen that was hiding the center of the rod: one solid, complete rod, or two rod pieces with a space where the screen had been. The authors reasoned that the infants would look longer at the test display that they perceived to be most discrepant from the center-occluded display.

The results showed that when the rod pieces were shown stationary behind the screen, the infants looked about equally at the two test displays, suggesting that they did not have a clear interpretation of the display. However, when the visible rod pieces moved as though they were a single object—moving side to side behind the screen with the same direction and speed—the infants showed a statistically significant preference for the broken over the complete rod—suggesting that infants expected to see a complete rod behind the screen. This pattern of results indicates that common motion allowed infants to form a clear interpretation of the display as composed of a single rod behind the screen. The authors also suggested that this ability to organize the visual world was not acquired by the infants but may rather be based on infants' innate conception of what an object is.

This claim was refuted when, in 1991, Alan Slater and his colleagues showed that when newborn infants are tested in a situation very much like the one previously described, they show the opposite preference for the test displays—looking reliably longer at the complete than at the broken rod, suggesting that they saw the partly occluded rod as composed of two separate pieces rather than as one single piece. These findings suggest that between birth and four months of age, infants learn to use common motion of object parts as a cue for uniting the parts of a partially occluded object. Thus, perceptual learning of object features and object motions may take place very early in life to help infants make sense of the world around them.

Encouraged by these findings showing effects of early learning and/or development, researchers went on to investigate what kind of information infants use when interpreting a visual display. For instance, Scott Johnson and Richard Aslin in 1995 showed that for infants between birth and four months old, small changes to the display such as (a) narrowing the occluding screen or (b) creating small openings in the screen that would allow infants to see glimpses of the rod through the screen helped these infants perceive the rod as connected behind the screen when it engaged in common motion. These findings show that infants' information processing skills are limited and can lead them to fail to see the connection between the visible rod pieces. But if infants have more information about the hidden object (provided by the narrower screen or the openings in the screen), they can overcome these limitations.

Parsing Adjacent Objects

Researchers have also been interested in assessing infants' perception of displays that offer a different kind of challenge—rather than having to infer the existence of a portion of an object that is not directly visible, when parsing adjacent objects (as in our blender and toaster example at the beginning of the entry), infants must determine where one object ends and another begins. What we know about infants' parsing of adjacent objects parallels what we know about their perception of partly occluded objects.

These studies have used a different method, based on the assumption that, if infants form a clear interpretation of a display, they should look reliably longer when that display is moved in a manner inconsistent with their interpretation than when the display is moved in a manner that is consistent with their interpretation. For instance, if you see a coffee mug sitting next to a paperweight, you would likely see the coffee mug as separate from the paperweight. So, if you were to reach out and pick up the coffee mug and the paperweight moved along with the mug, you would probably be surprised and look at this event longer than if the paperweight remained stationary when the mug was picked up. This is the basis on which these next several experiments were conducted.

A series of studies conducted by Amy Needham and colleagues has shown that by 8.5 months of age, infants perceive two objects that are different in shape, color, pattern, and texture as separate from each other, even when they are sitting against each other so that their sides are touching. Infants as young as four months of age can determine likely separation points between objects just on the basis of object shape, as long as the shapes of the objects in the display are relatively simple (i.e., regular, geometric shapes). Interestingly, shape is an important and reliable cue for adults' parsing of objects as well, and both of these facts (early use and consistent use even in adulthood) are linked to the fact that shape is a highly reliable cue for the locations of object boundaries. Shape is, of course, one of the most fundamental defining characteristics of an object: It is what differentiates a screwdriver from a hammer. Both are tools of about the same size, but they are used for different purposes, and their shapes would guide your selection of the hammer for driving a nail and the screwdriver for driving a screw.

What are the mechanisms that underlie infants' learning about how visible characteristics of objects can be used to predict separations between objects? One proposed mechanism is statistical learning, which holds that observers (adults or infants) are able to keep track of the co-occurrences of particular stimuli or stimulus parts over multiple views and extract the regularities that they experience over time. Based on past experiences with similar objects, the observer can then make predictions about the likely properties of new objects. Richard Aslin, Scott Johnson, and their colleagues have proposed that statistical learning (in addition to providing an excellent basis for learning many characteristics of language) can also provide observers with clear expectations about which visual shapes "go together." Using sequences of objects, they have shown that infants as young as two months of age quickly acquire expectations about which pairs of objects that appear in sequence are always adjacent to each other in the sequence (and which are not). Likewise, infants have expectations about which objects consistently appear near each other in a spatially defined grid (and which do not).

Together, these findings make clear that infants learn general principles (e.g., similar-looking surfaces are likely to belong to the same object, different-looking surfaces are likely to belong to different objects) that allow them to parse unknown objects into separate units. However, there is also evidence that infants learn about the exceptions to these rules—objects that may look like they are separate objects but nevertheless are part of the same unit.

One example of this kind of object is a key ring. Key rings (whether they are toy key ring rattles or actual metal keys) are composed of parts of different shapes, colors, and textures, and yet our experience shows us that the ring and keys of key rings are part of a single unit that moves as a whole. Thus, if infants are using similarity among these parts as the criterion for parsing this display, they should see the keys and ring (and possibly each of the keys) as separate objects. In contrast, if infants recognize the key ring from their prior experiences with it outside the lab, they should group the keys and ring as a single object that would move together.

How infants perceive key rings and similar objects was studied by Needham and colleagues in 2007. Using the expectation-based method previously described, this study showed that, by 8.5 months of age, infants expect the keys and ring to move together as a single object. Interestingly, this percept was altered by the systematic removal of distinctive features of the key ring (such as the cures that mark the edges of each key and the keys' action ends). Further, younger infants (just about a month younger) did not see the key ring as a single unit, but instead they saw the keys and ring as belonging to two separate objects. This further highlights the importance of experience and learning in infants' object perception.

Implications for Infants' Perceptual Development

The research described in this entry makes clear that although infants' visual and cognitive abilities are still developing, they have reasonably clear interpretations of the objects around them. On the basis of general principles, such as common motion and similarity, four-month-old infants can group the visible portions of a partly occluded object into a single unit and they can determine where one object ends and another begins. Infants also learn about certain kinds of objects that do not adhere to the typical principles of similarity, such as key rings, and respond to these displays as though they have learned about exceptions to these general rules. Together, both of these learning mechanisms could help infants make educated guesses about the organization of the world around them into separate objects.

Amy Needham

See also Object Perception; Perceptual Development: Visual Object Permanence and Identity; Perceptual Organization: Vision; Perceptual Segregation

Further Readings

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences USA, 99,* 15822–15826.

Johnson, S. P., & Aslin, R. N. (1995). Perception of object unity in 2-month-old infants. *Developmental Psychology, 31,* 739–745.

Needham, A., Cantlon, J. F., & Ormsbee Holley, S. M. (2006). Infants' use of category knowledge and object attributes when segregating objects at 8.5 months of age. *Cognitive Psychology, 53,* 345–360.

Slater, A., Morison, V., Somers, M., Mattock, A., Brown, E., & Taylor, D. (1990). Newborn and older infants' perception of partly occluded objects. *Infant Behavior & Development, 13,* 33–49.

PERCEPTUAL DEVELOPMENT: SPEECH PERCEPTION

Making sense of language hinges upon a skill called speech perception. Was that a bat, a hat, a rat, or a fat cat that sat? Small sound changes make for large differences in interpretation. What sound differences do infants perceive? When do they first recognize their native language? How do they discover words in the stream of speech, and how do their perceptual abilities change with time and exposure to that speech? These are the questions asked by researchers who study the development of speech perception.

Thanks to their work, it is now known that infants distinguish extremely subtle acoustic distinctions and recognize their native language from birth. By their second birthday, infants are experts at finding words in the fluent stream of speech and focus only on those acoustic differences that matter for their native language. This entry looks at two key areas in the development of speech perception: changes in infants' abilities to hear specific language sounds (called phonemes) and changes in their ability to discover words in the stream of speech (a task called segmentation).

Development of Phonological Perception

From birth, infants appear to hear speech differences in a categorical manner. They hear many, if not all, of the sounds of the world's languages and appear to notice certain critical differences from the earliest ages tested. For example, the acoustic difference between the /b/ in *big* and the /p/ in *pig* is subtle. The principle change has to do with the relationship between when the sound begins and when the vocal cords vibrate. In /b/ the vocal cords and sound begin simultaneously, whereas in /p/ the vocal cords begin their vibration noticeably later. This time difference is called the voice onset time (VOT). In the 1960s, Peter Eimas and colleagues systematically varied VOT to demonstrate that infants pay more attention to the difference between /b/ and /p/ rather than a similar-sized VOT difference between one /p/ and another /p/ (this time said with even more time between the onset of the sound and when the vocal folds vibrate).

How did they show this? They took advantage of something all infants do: suck. Using a method called the sucking procedure, these researchers gave infants an artificial nipple attached to a device that registered each time the infant sucked. In this case, the researchers played a "pa" over and over for English-learning infants. As the infants succumbed to boredom, their sucking decreased. When the audio was switched to "ba," their sucking increased, suggesting infants noticed the difference between the two sounds. Infant's sucking did not recover if the /p/ with a larger VOT was played, even though the size of the acoustic difference was the same. Patricia Kuhl subsequently demonstrated that even chinchillas make this distinction, suggesting that categorical perception of speech is not specific to humans and may relate to hardwired abilities of the nervous system to hear certain acoustic changes over others.

Infants even notice sound distinctions not found in their native language. For example, Japanese speaking adults have a hard time hearing the difference between /r/ and /l/, which leads them to confuse "surprise" with "supplies." Yet Japanese infants can hear the difference, and a wealth of research indicates this is true for a wide range of difficult-to-hear contrasts across a number of languages. This ability to distinguish difficult, non-native contrasts disappears by approximately 12 months. Such results suggest that though children can still learn new language sounds, around one year of age they have begun to specialize in their native language and have begun to ignore distinctions that are not important for that language. Kuhl called this focusing on the important sounds the *perceptual magnet effect.* That is, the native sound acts like a magnet: pulling similar non-native sounds to it (making them being heard the same).

Infants can hear all the sounds of the world's languages from the youngest ages tested, some sounds are easier for them to hear than others, and over the first year of life, infants begin to specialize in hearing just those distinctions used by their native language.

Development of Speech Stream Segmentation

Another of the tasks faced by infants as they perceive speech is discovering the individual words in the fluent flowing steam of speech. This isn't as easy as it might first appear. Words in speech are not separated by spaces, commas, and periods, as in the printed word. Furthermore, words in isolation are actually relatively rare in the input to infants, yet infants must pull out some words from passages if they are ever to discover what words mean.

Can infants find even one word in the acoustic speech stream? In 1995, Jusczyk and Aslin used a procedure called the head turn preference procedure to demonstrate that 7.5-month-olds can find one word in a speech stream. In this procedure, infants sat on their parent's lap in the center of a three-sided booth with lights on each side. The lights flashed as a passage written around a target word was played. For example, infants heard: "The cup was bright and shiny. The clown drank from the red cup. His cup was filled with milk . . ." After infants had listened to the passage for a set amount of time, the researchers played the target word from this passage in isolation (e.g., "cup, cup, cup") versus a word the infants hadn't heard (e.g. "feet, feet, feet"). Infants looked longer at the lights when the word "cup" was played indicating they had successfully discovered that word in the passage. Infants also looked longer when "cup" was played versus a similar word, "cut," indicating their acoustic memory was quite detailed.

The head turn preference procedure has since been used by hundreds of labs to test the boundaries of infant speech segmentation. For example, breaks in the acoustic stream often do not correlate exactly with words. Thus, in the phrase "my speech" there is silence after the /s/ and before the /p/. If infants segmented at the break, they would hear "mice peach." In fact, before their first birthday, infants segment such confusable phrases correctly, suggesting a remarkable knowledge about the sound patterns of speech. It is now known that infants correctly segment speech using statistical cues, allophonic cues, and stress cues.

The first way that infants seem to discover words in a fluent stream of speech is through the use of statistics. That is, infants notice when the same sounds and syllables tend to appear together. So in the previous study, infants appeared to notice that /cu/ and /p/ go together. This is especially true for familiar words, and Jennifer Saffran, Richard Aslin, and Elissa Newport have shown infants can even do this with limited exposure to a made-up language in which the "words" consist of three syllables that always go together. Thus, these researchers played a continuous "passage" in which /pi/, /go/, and /la/ always went together: "**pigola**bikuti**pigola**pabiku**pigola**." Note that although "pigola" is printed in bold, the actual passages had no clues to indicate the correct word. Nonetheless, infants were able to discover the word in the passage. This suggests that when adults use the same word multiple times in a phrase (e.g., "Do you see the ball?" "Look at the ball." "It is a ball!") infants recognize that /ba/ and /l/ always go together, and use this information to discover the word.

Such work suggests that the first words infants discover in speech are the most frequent ones. This appears to be especially true for one word infants hear quite a bit—their own name. Not only can infants recognize their own name as early as six months, but recent work indicates infants can use their name to discover additional words. So for example, if Julien, the infant, hears "Julien's cup was bright and shiny. A clown drank from Julien's cup . . ." then even at six months, Julien could pick out the word, cup—a full month and a half earlier than if the name was not present.

Also, over the first year of life, infants can learn to recognize where certain sounds typically appear (called *phonotactics*), and the particular ways those sounds are said (called *allophonics*). For example, the way one says a /p/ at the beginning of a word is different from how one says it at the end of a word. Likewise, in English, the sound /ng/ never appears at the beginning of a word, whereas a special kind of /l/ (called a "light" /l/), never appears at the end of a word. Infants can use this information about how phonemes differ, to make educated guesses about where new words are likely to segment, thereby distinguishing "my speech" from "mice peach."

Research has also shown that infants have an easier time discovering a word if it is at the end or beginning of a phrase, and that infants are more successful if the speaker is using an exaggerated type of speech known as *infant-directed speech*. That is, adults the world over tend to adopt a higher register of speaking when talking to infants and tend to become more exaggerated and extreme in the sounds they use. These tendencies apparently make the words and important features more distinct in the stream, which makes it easier for the infants to find the words. For example, one exaggerated feature is something called *lexical stress,* where words are accented. Many languages tend to use a particular stress pattern when words begin. So for example, English tends to accent the first syllable of words (e.g., BARbie, SCIence, DOCtor). If infants "knew" this, then they could simply assume a new word begins whenever they hear a stressed syllable. Not only do English-learning infants prefer words with the stress on the first syllable, but by 7.5 months of age, they appear to adopt exactly this type of strategy, called the metrical segmentation strategy, segmenting words at the accented syllable, although their use of this strategy appears to depend on statistics.

Babies' lives are often anything but quiet. Whether from parents talking, siblings shouting, TV or radios blaring, or even those "soothing sounds" CDs, children frequently hear speech under distracting circumstances. How well do infants segment speech under such situations of noise and distraction? The answer is that infants don't deal with distraction very well. Using the head turn procedure, George Hollich, Rochelle Newman, and Peter Jusczyk found that if someone else is talking at the same volume or higher, seven-month-olds are unable to pull out the individual words. However, if infants could see a video of the person speaking, infants can pick out individual words. This suggests the importance of visual information in speech perception and stream segmentation. Interestingly, infants were able to segment the speech stream even if a wiggly line that matched the woman's voice (not unlike HAL in 2001 or KIT from Knight Rider) was played. So infants likely can use any simultaneous movement to help them hear better—even gestures or head movements.

Infants come to segment the fluent stream of speech using a combination of statistics, location in the phrase, and the stress on the words, all of which tend to be exaggerated by adults in the form of infant-directed speech and that can be heard better when the sound is combined with visual information.

George Hollich

See also Amodal Perception; Bayesian Approach; Computer Speech Perception; Infant Perception; Infant Perception: Methods of Testing; Language; Perceptual Development: Hearing; Perceptual Development:

Intermodal Perception; Speech Perception; Word Recognition

Further Readings

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech steam segmentation. *Psychological Science, 4,* 298–304.

Hollich, G. (2006). Combining techniques to reveal emergent effects in infants' segmentation, word learning, and grammar. *Language & Speech, 49,* 3–19.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5,* 831–843.

Jusczyk, P. W. (1999) Narrowing the distance to language: One step at a time. *Journal of Communication Disorders, 32,* 207–222.

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science, 12,* 110–114.

Werker, J. F., & Tees, R. C. (1999). Experiential influences on infant speech processing: Toward a new synthesis. In J. T. Spence (Ed.), *Annual review of psychology* (Vol. 50, pp. 509–535). Palo Alto, CA: Annual Reviews.

PERCEPTUAL DEVELOPMENT: TASTE AND OLFACTION

Unlike senses such as vision and audition, where information is transmitted through light and sound waves, our senses of smell and taste are mediated by molecules that stimulate receptors throughout the nasal and oral cavities, respectively. As a result, smell and taste are commonly referred to as *chemical senses,* and in combination with other inputs, such as texture and temperature, are responsible for determining the flavor of the foods and beverages we ingest. Our ability to perceive and integrate these sensory stimuli plays an essential role in determining our food consumption by attracting us to potential food items and warning us of the presence of potentially harmful substances. Presumably due to their adaptive value, these perceptual systems are functional at birth. However, despite their early emergence, the chemosensory systems continue to mature postnatally and are influenced by early flavor experiences. Thus, as discussed in this entry, the flavor world of infants and children is very different from that of adults.

Development of Taste Perception

Taste buds, which contain clusters of receptors that will ultimately be sensitive to the basic tastes of sweet, salty, sour, bitter, and umami, begin to emerge with the formation of the mouth and palate just eight weeks after conception. By 13 weeks, the taste buds begin to resemble those of the adult and are found throughout the oral cavity. It is probably not a coincidence that the timing of these developmental changes corresponds to the emergence of fetal sucking and swallowing behavior. By term, infants are actively swallowing between 500 to 1,000 milliliters (ml) per day of amniotic fluid. This not only represents a major route of amniotic fluid absorption, but it also serves to stimulate the taste buds and influence their synaptic connections. Because the chemosensory composition of the amniotic fluid is in constant flux throughout pregnancy, especially once the fetus begins to urinate, the newly developed taste buds are exposed to varying concentrations of a variety of chemosensory stimuli, including sugars, sodium and potassium salts, and various acids. Behavioral studies that have tested preterm infants between 33 and 40 weeks postconception have shown that they produce more frequent and stronger sucking responses to a sucrose-sweetened nipple than to an untreated nipple. Thus by the last trimester, taste buds are capable of detecting sweet tastes and communicating information to structures that are responsible for organizing and controlling affective behaviors within the central nervous system.

More extensive testing of full-term newborn infants has revealed that they not only have heightened preferences for sweet tastes, but as indicated by consumption tests, they are also able to discriminate between various concentrations of dilute sucrose solutions and between different types of sugars. Additionally, as originally described by Jacob Steiner in 1973 and replicated by others, within two hours after birth infants express positive affective reactions, such as facial relaxation, sucking responses, and lip retractions that resemble smiles, in response to a drop of sweet solution placed on their tongue, which are interpreted as hedonic (i.e., liking) responses. Although individual differences in children's sweet taste preferences exist largely due to the interplay between genetics and experience, heightened preferences for sweet taste have been observed in children throughout the world. These findings suggest that infants have an innate preference for sweet tastes. This biological predisposition may have evolved to attract infants to sources of calories such as breast milk, which is predominantly sweet.

In addition to being sweet, breast milk is high in glutamate content, which imparts a savory or umami taste to foods such as meat, fish, cheese, and vegetables, and is commonly used as a food additive in the form of monosodium glutamate (MSG). Although infants do not appear to respond to aqueous solutions of MSG, they display distinct facial expressions indicative of liking in response to a soup that contains MSG relative to soup alone. Long before their first taste of solid food, newborns display strong preferences for the basic tastes of sweet and umami.

In marked contrast, when newborn infants are exposed to concentrated bitter taste stimuli at only two hours of age, they respond with facial expressions of distaste, which include specific actions that block swallowing such as gaping. In addition to blocking ingestion, this visually striking orofacial response is readily identified by adults, and thus provides the preverbal infant with an efficient means of communicating their distaste for the bitter substance. According to some theories, such stereotypical responses may have evolved as a defense mechanism to protect infants from ingesting poisons and toxins, which often taste bitter. Interestingly, although concentrated bitter solutions elicit strong orofacial responses soon after birth, infants appear to be relatively indifferent to weak to moderate concentrations of bitter until about two weeks after birth as measured by consumption tests, suggesting that their sensitivity to bitter tastes continues to develop postnatally.

Although less scientific research has been conducted on the ontogeny of sour taste, existing studies suggest that this basic taste also elicits facial expressions of distaste (i.e., gaping and lip pursing) and reduced consumption in newborn infants. However, for many children, this initial dislike is temporary and gives way to a heightened preference for extreme sour tastes sometime before the age of five years of age. In a recent study, more than one-third of five- to nine-year-old children, but none of their mothers, preferred high levels of a sour taste presented in gelatin. These individual differences in children's sour preferences appeared to be related to their willingness to try unknown foods and their early experience with sour-tasting foods and formulas (such as protein hydrolysate formulas, which have an unpalatable taste that consists of both sour and bitter notes). Of course, these experiential effects do not rule out the possibility that maturational changes in taste perception may underlie the heightened sour preferences observed in childhood.

Unlike the other basic tastes, salt does not appear to reliably elicit either distinctive facial expressions or differential intake in newborn infants, suggesting that human newborns may be indifferent to this basic taste. Consistent with developmental shifts in salt perception in other mammalian species, infants' preference for salt emerges at approximately four months of age. As suggested in animal model studies, the shift from indifference to preference appears to be largely due to maturation of neural mechanisms underlying taste perception. Similar to that observed for sweet taste, children's heightened preference for salt continues until late adolescence. In addition to these maturational changes, research suggests that perinatal experiences may also play an important role in the development of children's salt preferences later in life. Infants whose mothers reported moderate to severe morning sickness, which produces dehydration and salt depletion, had higher salt preferences as demonstrated by facial expressions and intake.

Development of Olfactory and Flavor Perception

Relative to taste, less is known about infant's olfactory perception and preferences. For example, although infants appear to have innate likes and dislikes for various tastes, it is currently not clear whether there are inherent differences in their hedonic responses to odors at birth. Researchers do know, however, that newborn infants can discriminate a wide variety of odors and encode and retain olfactory memories. As discussed in the following paragraphs, these abilities develop well before birth.

By the 11th week of gestation, just a couple of weeks before the taste buds have reached maturity, the olfactory bulbs and receptor cells that are needed to detect olfactory stimuli attain adultlike morphology. Despite their early development, olfactory neurons are not capable of detecting odors until the last trimester (approximately the 28th gestational week). By this point, the human fetus is actively inhaling more than twice the volume of amniotic fluid that it swallows. Because the odor of the amniotic fluid is affected by a variety of factors, including the mother's unique odor signature and her dietary choices, the continual flow of this complex fluid through the nasal passages provides the fetus with a variety of unique experiences. The fetus not only perceives these olfactory changes, but as discussed next, there is evidence that they are encoded and remembered.

Perhaps because of its familiarity, infants appear to be soothed by the odor of amniotic fluid after birth. Those who are separated from their mothers cried significantly less when exposed to the odor of amniotic fluid relative to the odor of the mothers' breast. Moreover, newborns reliably orient toward a breast or gauze pad that is treated with a small amount of amniotic fluid relative to an untreated control. This initial preference for amniotic fluid is fairly transient, however, and over the first couple of days of life it gradually dissipates as breast odors become increasingly more preferred. Research has shown that by about the first week of life, both breastfed and formula-fed neonates prefer breast odors when compared to axillary odors of the same women. This general attraction to odors emanating from the breast, which is likely facilitated by the overlapping odor profiles of amniotic fluid and breast milk, plays an important role in guiding the infant to the nipple area and assisting nipple attachment and breastfeeding.

In addition to facilitating feeding behavior, breast odors reflect the unique "odor signature" of the mother and, as a result, provide the basis for maternal recognition by the neonate. By three days of age, infants prefer their own mother's breast odors relative to those of another woman. The ability of infants to discriminate the odors of their mothers from those of other lactating women is not limited to the recognition of breast odors. Breastfed infants can also discriminate natural odors originating from their own mother's underarms and neck and they learn to prefer artificial scents, such as perfumes, that are worn by their mother. That formula-fed infants do not display similar recognition of their mothers' axillary odors implies that this early recognition may be a function of postnatal learning processes. Because infants have their nostrils in close proximity to their mother's skin when they breastfeed, they may learn to associate their mother's odor with reinforcing stimuli such as warmth and food intake. Once the breastfed infant becomes familiar with the unique scent of the mothers' breasts, they are capable of recognizing that same scent when it is produced by other parts of the body.

The infant's capacity to learn about odors is not limited to its mother's genetically determined olfactory signature, as previously mentioned; they are also capable of learning about her dietary preferences. Because odor volatiles from the mother's diet are transmitted to amniotic fluid, infants begin to learn about the foods within their culture before birth. This was demonstrated in a study that reported that newborn infants preferentially orient toward the odor of garlic if their mothers ate garlic during pregnancy, whereas infants whose mothers avoided consumption of garlic oriented toward an unscented swab. If the mother chooses to breastfeed, these olfactory experiences continue after birth, because like amniotic fluid, the flavor of her milk reflects her dietary choices. To date, psychophysical studies have revealed that a wide range of odor volatiles that are either ingested (e.g., alcohol, garlic, vanilla, carrot) or inhaled (e.g., cigarette smoke) by the mother are transmitted to her breast milk. These early olfactory experiences may provide a mechanism for the development of flavor preferences that are similar to those of the mother. This was demonstrated in a study in which infants whose mothers consumed carrot juice either during pregnancy or the first two months of lactation displayed fewer negative responses while eating carrot-flavored cereal than plain cereal at six months of age. No such preference for the carrot flavor was observed in a third group of infants whose mothers did not drink carrot juice during pregnancy or lactation. Although more research is needed to understand how these early flavor experiences affect later food preferences, repeated exposure to a variety of flavors within amniotic fluid and breast milk may be one of the first ways that children learn about the foods within their culture and serve to enhance their acceptance and preference for novel foods at weaning.

Catherine A. Forestell

See also Experience-Dependent Plasticity; Flavor; Infant Perception: Methods of Testing; Olfaction; Olfaction and Reproductive Behavior; Taste and Food Preferences

Further Readings

Forestell, C. A., & Mennella, J. A. (2008). Food, folklore and flavor preference development. In C. Lammi-Keefe, S. C. Couch, & E. Philipson (Eds.), *Handbook of nutrition and pregnancy* (pp. 55–64). Totowa, NJ: Humana Press.

Ganchrow, J. R., & Mennella, J. A. (2003). The ontogeny of human flavor perception. In R. L. Doty (Ed.), *Handbook of olfaction and gustation* (2nd ed., pp. 823–846). New York: Marcel Dekker.

Liem, D. G., & Mennella, J. A. (2003). Heightened sour preferences during childhood. *Chemical Senses, 28,* 173–180.

Mennella, J. A., Jagnow, C. P., & Beauchamp, G. K. (2001). Prenatal and postnatal flavor learning by human infants. *Pediatrics, 107,* e88.

Porter, R. H., & Winberg, J. (1999). Unique salience of maternal breast odors for newborn infants. *Neuroscience and Biobehavioral Reviews, 23,* 439–449.

Rosenstein, D., & Oster, H. (1988). Differential facial responses to four basic tastes in newborns. *Child Development, 59,* 1555–1568.

Steiner, J. E. (1977). Facial expressions of the neonate infant indicating the hedonics of food-related chemical stimuli. In J. M. Weiffenbach (Ed.), *Taste and development: The genesis of sweet preference.* Washington, DC: U.S. Government Printing Office.

PERCEPTUAL DEVELOPMENT: TOUCH AND PAIN

This entry describes how touch and temperature are perceived in neonates and the behavioral and physiological responses of neonates to pain.

Touch Perception

Touch stimuli cause changes in the skin that give us the sensations of pressure, warmth, and vibration. Touch is the earliest sensory system to develop in all species. When a human embryo is less than an inch long and less than two months old, the skin is already highly developed. When the palm is touched at two months gestation, the fingers grasp the palm. The fingers and thumb will close at three months when the palm is touched. The fetus can turn toward a tactile stimulus, much like a rooting reflex, and the fetus responds to electrical stimuli and puffs of air that are even difficult for adults to discriminate. These very sophisticated perceptual skills occur as early as three months gestation.

Touch Perception in the Neonate

Research suggests that touch discrimination by mouth and by hands occurs as early as the newborn period. Different texture nipples (nubby versus smooth) can be discriminated by the newborn's mouth and by their hands. Using a habituation/recovery paradigm, the newborn explored the different texture nipples by their mouths and by their hands. After experience with one nipple, they habituated or stopped sucking. When the new texture nipple was presented, they started to suck again, indicating that they discriminated between the first and second nipple.

Texture perception has also been investigated in three-day-old infants using a habituation paradigm with a smooth or granular object. After holding the object, babies were given either a familiar or a new textured object. Two dependent measures were recorded: (1) holding time was used to assess habituation as well as reaction to novelty, and (2) hand pressure frequency exerted on the object was used to examine the neonates' ability to adjust their manipulation to the texture of new objects. Both measures revealed perception of the texture of objects.

Weight perception has also been studied in young infants. Light-weight (two grams) or heavier-weight (eight grams) objects (vials of cotton or pellets, respectively) were placed in the right hand of full-term newborns of depressed or nondepressed mothers. After the infants habituated one weight by hand, they were tested with the opposite weight object. The infants of the depressed mothers did not respond to the novel weight, and only 15% of those infants showed hand movements that might have facilitated their perception of the object's weight (e.g., hand-to-mouth or hand-to-face, turning/moving of the wrist or hand). In contrast, 78% of the infants of nondepressed mothers showed hand activity that would lead to weight perception, and, as a group, they held the novel weight longer, suggesting that they had perceived the weight change, whereas the infants of depressed mothers failed to make that discrimination.

Another group of investigators studied weight perception in older (one-year-old) infants. The authors not only documented weight perception by one-year-old infants but also the ability to adjust their hand manipulation of the weights. Two dependent measures were recorded: (1) holding times in order to assess habituation as well as reaction to novelty, and (2) manual pressure exerted on the object to investigate the infants' ability to adjust their manipulation to the object's weight. Both measures revealed infants' haptic perception of weight, but the second measure suggested that the one-year-olds had developed manual manipulation as another indicator of their weight perception.

Cross-modal or intersensory perception—-perception across different sensory modalities—also occurs at the newborn period. In a demonstration of this skill, newborns were given different shapes to handle and then they were shown a picture of different shape objects. The infants looked longer at the object they had handled, suggesting that newborns can visually recognize the shape of an object that they have previously manipulated with their right hand when it's out of sight. Newborns can extract shape information in a tactual format and transform it into a visual format before they have had the opportunity to learn from the pairings of visual and tactual experience. However, the same authors later showed that this ability was absent when the left hand was involved.

Temperature Perception

The related sense of temperature has not been studied with newborns until recently. In a study on temperature perception by mouth, researchers gave neonates of depressed and nondepressed mothers cold and warm nipples on alternating trials. The newborns of depressed mothers sucked twice as long as those of nondepressed mothers, suggesting that they were aroused, dysregulated, or overactive. The newborns did not show a preference for cold or warm nipples, but the neonates who received the cold nipple on the first trial sucked significantly more on the subsequent two trials, suggesting that for some reason the cold nipple led to more sucking, perhaps to warm the nipple.

In a study on temperature perception by hand, researchers gave newborns of depressed and nondepressed mothers temperature tubes containing cold or warm water to hold onto with their hands. Both groups of infants habituated the warm and cold tubes, as indicated by a decrease in holding, and they showed dishabituation, as indexed by an increase in holding the novel temperature tube. The newborns of depressed mothers, however, took twice as much time to habituate, possibly because they did not actively explore the tubes with their hands.

Pain Perception

Painful stimuli lead to a stress response as early as the newborn period. Preterm infants experience up to 60 invasive procedures before being discharged from the neonatal intensive care unit. The heel stick has been the most common procedure. Animal and human research have demonstrated that repetitive pain has adverse effects on the neurodevelopment of the neonate. Although there are still some neonates who do not receive analgesia during invasive procedures because they are not thought to experience pain, neonates clearly show physiological and behavioral responses during heel sticks and other invasive procedures.

Behavioral Responses

Preterm infants typically show behavioral distress during invasive procedures. Although facial expressions and crying are the most widely studied responses to pain, body movements have also been reported as distress behaviors in preterm infants experiencing heel sticks. Extension of the arms and legs (80%) and finger splays (70%) are examples noted in a heel stick study. In that study, preterm infants who experienced invasive procedures more frequently and were younger gestational age showed more motor stress behaviors after the heel stick. The flexion reflex or withdrawal of the limb from the stimulus has also been reported as a response to invasive procedures. A study also used photogrammetry to document the flexion reflex during routine heel sticks. In the same study, the infants withdrew both legs and cried immediately after the first heel stick. This motor response is similar to the adult's response to pain.

Typical facial expressions following an invasive procedure are a brow bulge, eye squeeze, lip purse, stretched mouth (horizontal or vertical), and chin quiver. Male and female infants respond differently to pain. In a study on cry sounds and cry faces in response to pain, gender differences were noted. Female infants cried louder than males, and their cries were higher-pitched than those of males. Males and females did not differ on facial expressions in response to pain.

In one study, infants who had the most invasive procedures since birth showed fewer facial reactions to heel sticks, suggesting that the infants might have habituated painful stimuli. However, another study found no relation between the number of invasive procedures experienced by the preterm infant and their behavior during heel sticks. In a third study, experienced infants appeared to anticipate the painful stimulus, but no change in reactivity was noted over days for most preterm infants observed, suggesting that the infants never grew accustomed to the heel sticks. These mixed data may relate to the timing of data collection. Some researchers conducted observations immediately after the heel stick (30 and 60 seconds after), and others waited as long as two minutes after the heel stick to code behavior. The infants' negative reactions typically occurred immediately after the heel stick, suggesting a reflexive response.

Infant crying may be the most obvious pain response to invasive procedures. However, some studies suggest that some preterm infants do not cry, perhaps because of their weakness or shutting down responses. In a longitudinal study on preterm infants, responses to repeated heel sticks were most often facial expressions, as described in the previous paragraph.

Physiological Responses

Pain has been measured not only by facial expressions but also by physiological measures, such as heart rate. A study that involved a routine blood test by heel stick showed both facial grimacing and increased heart rate. In another heel stick study, facial expressions, behavioral states, and fussing were recorded. Here, an age difference was noted, with younger infants displaying fewer of these responses and spending more time sleeping. Going into deep sleep following pain, for example after a circumcision, is called the conservation withdrawal response. Stronger behavioral responses were accompanied by stronger heart rate responses to the heel stick.

Disorganized physiological responses have also been noted for neonates during and following invasive procedures, including changes in heart rate, respiratory rate, and blood pressure. For example, in studies on preterm and full-term neonates experiencing heel sticks and following circumcision, increases in heart rate and blood pressure have been noted both during and after the procedures. And, more complex responses have been reported, including increased heart rate, increased respiratory rate, increased systolic blood pressure, and unchanged diastolic blood pressure.

Heart rate elevations, as high as 18% above baseline, have been reported for preterm infants during heel sticks. The physiological disorganization in response to repetitive heel sticks may redirect energy and oxygen, may disturb sleep cycles, and increase morbidity and mortality.

Greater experience with painful procedures has led to a downregulation of cortisol responses to subsequent neonatal intensive care unit (NICU) stressors. Cortisol is assayed from saliva, which can be collected by a new filter paper method that simply involves touching the infant's tongue with the filter paper for a couple of seconds. Infants who received numerous invasive procedures at birth had elevated cortisol levels at eight months of age. Repeated saliva samples collected from preterm infants suggested the early development of circadian-like patterns.

In an immune study on pain in premature infants, the pain led to increased cortisol and also decreased immune cells. Animal research indicates that toxic chemicals are also released during repetitive painful events in the neonatal rat, which may have deleterious effects on the developing central nervous system (CNS). In turn, these negative effects on the CNS may affect pain thresholds and long-term behaviors. Thus, decreasing or preventing the disorganized responses to invasive procedures is important for infant well-being and normal development.

These data suggest that newborns, and even preterm newborns, perceive pain. It is not clear, however, whether those perceptions change over time or how they differ from pain perception by adults. Perceiving pain and being able to cry for help from the first days of life seems highly adaptive.

Tiffany Field

See also Pain: Assessment and Measurement

Further Readings

Anand, K., Johnston, C., Oberlander, T., Taddio, A., Tutag Lehr, V., & Walco, G. (2005). Analgesia and local anesthesia during invasive procedures in the neonate. *Clinical Therapeutics, 27,* 844–876.

Field, T., & Goldson, E. (1984). Pacifying effects of nonnutritive sucking on term and preterm neonates during heelstick procedures. *Pediatrics, 74,* 1012–1015.

Goubet, N., Clifton, R., & Shah, B. (2001). Learning about pain in preterm newborns. *Journal of Developmental and Behavioral Pediatrics, 22,* 418–424.

Grunau, R. E., Holsti, L., Haley, D. W., Oberlander, T. F., Weinberg, J., Solimano, A., et al. (2005). Neonatal procedural pain exposure predicts lower cortisol and behavioral reactivity in preterm infants in the NICU. *Pain, 113,* 293–300.

Hernandez-Reif, M., Field, T., & Diego, M. (2004). Differential sucking by neonates of depressed versus non-depressed mothers. *Infant Behavior and Development, 27,* 465–476.

Hernandez-Reif, M., Field, T., Diego, M., & Shay, L. (2001). Weight perception by newborns of depressed versus non-depressed mothers. *Infant Behavior and Development, 24,* 305–316.

Hernandez-Reif, M., Field, T., Diego, M., & Shay, L. (2003). Haptic habituation to temperature is slower in newborns of depressed mothers. *Infancy, 4,* 47–63.

Lhote, M., & Streri, A. (1998). Haptic memory and handedness in 2-month-old infants. *Laterality, 3,* 173–192.

Streri, A. (2003). Cross-modal recognition of shape from hand to eyes in human newborns. *Somatosensory and Motor Research, 20,* 13–18.

Streri, A., & Gentaz, E. (2003). Cross-modal recognition of shape from hand to eyes and handedness in human newborns. *Neuropsychologia, 42,* 1365–1369.

PERCEPTUAL DEVELOPMENT: VISUAL ACUITY

Visual acuity (VA) is a measure of spatial resolution of the visual system, which is determined at least in part by the ability of the eye to focus incoming rays of light directly onto the retina at the back of the eye. For mature, cooperative, verbal individuals, VA is typically measured by having the subject identify letters on an eye chart that is at a precise distance from the individual. The stimuli are black letters on a white background, yielding the highest possible contrast. In a clinical setting, VA is generally defined as a fraction with the denominator specifying the distance at which the subject sees what most individuals see at 20 feet. Normal VA of 20/20 can be achieved for most individuals naturally or by use of optical correction to focus the rays of light onto the back of the retina. The standard and familiar Snellen chart has been in use since the mid-1800s. The singular "E" at the top of this chart makes a good reference point: If this is the only letter that an individual can identify, his or her VA is approximately 20/200, or at 20 feet this individual can see what most people see at 200 feet. Typically, an individual with 20/200 or worse VA is considered "legally blind."

Still, the question remains: How do you know what an infant can see? True, these youngest of humans will not be reading an eye chart or even identifying visual shapes. But researchers have devised some relatively simple yet sophisticated ways to determine what infants see. Behavioral and electrophysiological measures are the two most commonly used techniques for evaluating visual acuity in human infants. These techniques are quite different from each other, but both have their place in research as well as clinical applications. This entry describes the behavioral and electrophysiological measures of VA and newborns, along with the underlying physiological limitations relating to VA and newborns.

Behavioral Measures

Researchers can learn about infant visual acuity by very carefully watching their looking behaviors under controlled conditions. When an infant is shown two visual stimuli presented side by side against a uniform background, he or she will typically look longer at one pattern than another. Researchers can use this looking behavior to infer that the infant can tell the difference between the two patterns. This strategy for evaluating infant visual perception was first pioneered in the early 1960s by Robert Fantz. His series of experiments were designed to determine if newborn human infants were able to perceive any visual patterns. Fantz used a paradigm known as preferential looking (PL). His strategy was to position an infant in front of a visual display that occluded the infant's view of anything but the screen in front of him. Two patterns were displayed on the screen, one to the right and one to the left of the center. An observer, positioned behind the screen with no view of what the infant was looking at, could watch the infant's eyes and record the amount of time that the infant looked at each pattern. The PL protocol was used to determine infants' preferences for a variety of patterns, such as black and white horizontal stripes versus a black-and-white bull's-eye. The original series of experiments focused on differentiating infants' preferences based on such dimensions as complexity, number, size, regularity, and curvature. One limitation of this paradigm is that group data are required to make any statements regarding visual perception. In other words, vision could not be quantified for an individual infant. Furthermore, the pairing of patterns and preference for one pattern over the other was not guided by a cohesive, reliable theory of visual information processing.

In the mid-1970s, Davida Y. Teller (and almost simultaneously Richard Held) determined that Fantz's PL methodology could be merged with more rigorous techniques of sensory psychophysics. The union of the two-alternative forced-choice (TAFC) paradigm, used to estimate sensory thresholds in cooperative human adults, with PL yielded a sound laboratory technique for evaluating visual information processing abilities in individual human infants: forced-choice preferential looking (FPL). In the TAFC paradigm, instead of being shown two separate patterns, an observer views a display with one pattern placed either to the left or right of center. The important feature is that the rest of the display has a uniform appearance and that, when the pattern is sufficiently small or dim, it becomes indiscriminable from the background. In other words, the display appears to be a homogenous field and the observer cannot see any difference between the right and left sides. For estimates of VA, the patterns are typically patches of black and white stripes called gratings. The background is gray and is carefully matched to the patterns with regard to overall brightness. This means that, as the width of the stripes become smaller and smaller for different stimulus presentations, the pattern becomes harder and harder to see against the gray background. Eventually, the pattern is so small that is it indiscriminable from the gray surround and the entire display appears uniform.

Fantz's PL experiments with infants had already established that infants will look at visual patterns. The FPL technique relies on the assumption that given a choice between a homogeneous field and a pattern, infants will look at the pattern. This assumption has proven to be exceptionally fruitful in allowing researchers to develop an image of the infant's visual world. The FPL technique requires an observer, who is masked to the infant's view, to use the infant's looking behavior to decide if the pattern is on the right or the left. The observer's judgments can be determined to be correct or incorrect based on the actual location of the pattern. If the stripes are wide enough, the infant will demonstrate very clear looking behaviors and the observer will be correct 90 to 100% of the time. At some point the stripes will be invisible to the infant and the observer will be correct only 50% of the time (chance). Typically, a threshold requires that at least five different patterns be presented 20 times each, or 100 trials total. Sensory threshold, or visual acuity, is specified as the width of stripe for which the observer is correct, approximately 75% of the time.

The FPL technique afforded the opportunity for quantitative assessments of sensory visual information processing in individual infants, and with it an entirely new realm of study began. So, what can a newborn human infant see? Using behavioral measures, newborn visual acuity is approximately 20/600. By 12 weeks of age, a typically developing infant will have approximately 20/200 acuity, and by 12 months of age, visual acuity is almost adultlike at 20/50. Behavioral measures generally do not show full adult acuity of 20/20 until somewhere around the age of 5 years.

Pediatric eye care physicians were most interested in being able to use quantitative measures of VA with their preverbal and nonverbal patients being treated for visual disorders. Still, after years of intensive laboratory work, investigators recognized that the FPL protocol required an intensive amount of time and would never really be applied for clinical purposes. At this point, Velma Dobson and Davida Y. Teller recognized that a more subjective interpretation of the infant's looking behaviors might yield a clinically useful test. The acuity card procedure has provided exactly that. In this procedure, the stimuli are presented to the infant in a "puppet stage." The observer is behind the screen and watches the infant through a peephole in the acuity card, which measures 25 by 55 centimeters and is very portable. Each of the 17 acuity cards has patches of black and white stripes on them that are different widths. Now the observer is allowed to use the quality of the infant's looking behaviors to make a judgment about whether or not the infant preferentially fixates the stripes. The observer is required to make a "Yes" or "No" decision for each acuity card: Can the infant see the stripes? This subtle difference in the protocol yielded a technique that could be completed in a very short amount of time, typically three to five minutes of testing for each acuity estimate. Fortunately, the estimates of VA from the acuity card procedure correspond very well with those obtained from the lengthier FPL procedure. The hours of laboratory work have resulted in a technique that has been used in many federally sponsored clinical trials and has yielded a tremendous amount of information for scientists as well as clinicians.

Electrophysiological Measures

Electrophysiological measures can also be used to estimate VA in young infants. In this procedure, brain waves are recorded from alert, attentive infants while they are watching patterns that change in a systematic manner on a video screen. These patterns are typically black-and-white checkerboards or stripes whose squares (or stripes) alternate back and forth from white to black or black to white simultaneously. That is the black stripes (or checks) become white and vice versa. These patterns are displayed for up to one minute to record the visually evoked potential (VEP). Different sizes of stripes or checks will yield different timing and amplitude of the response. This technique is similar to that used for other brain wave recordings (EEG) or heart monitoring (EKG). VAs measured from infants in this manner are typically higher than those measured behaviorally. Different laboratories have demonstrated VEP VAs that are adultlike by about 12 months of age in human infants. VEP techniques require expensive equipment and technicians with more sophisticated training, so although the VEP measures can be quite useful, the requisite setting is not as commonly available as behavioral testing setups.

Underlying Physiological Limitations

Most adults achieve good VA by focusing incoming rays of light directly onto the retina at the back of the eye, either naturally or with the aid of lenses. However, the ability to focus light onto the back of the eye is not the limiting factor for the newborn and developing human infant. There are numerous physiological changes that occur at all levels of the visual system as the infant matures. All of these changes will have an impact on the ability of the visual system to discern fine detail.

First, the shape and size of the eyeball itself increases with age from infancy to early childhood. The infant's smaller eye has more difficulty changing the shape of the lens and focusing the light onto the back of the eye.

Second, at the level of the retina, the photoreceptors are considerably different in shape and size. A newborn infant's photoreceptors are short and fat compared with those present in a mature visual system. The structure of the photoreceptors makes it more difficult for them to capture wavelengths of light and, thus, infant photoreceptors have low sensitivity to light. The shape of the photoreceptors also means that they are less densely packed in an infant retina and, hence, cannot discriminate the fine detail that an adult with more densely packed photoreceptors can.

Finally, cortical factors account for the majority of the changes in visual processing abilities. Initially, the cortex develops an "exuberance" of connections from birth to about 15 months of age. This period is followed by "pruning" of the cortical synapses that leads to a refinement in visual processing. At the same time, the neural fibers are becoming myelinated and are therefore transmitting signals faster and more efficiently.

E. Eugenie Hartmann

See also Visual Acuity

Further Readings

Dobson, V., & Teller, D. Y. (1978). Visual acuity in human infants: A review and comparison of behavioral and electrophysiological studies. *Vision Research, 18,* 1469–1483.

Hartmann, E. E. (1995). Infant visual development: An overview of studies using visual evoked potential measures from Harter to the present. *International Journal of Neuroscience, 80,* 203–235.

Hartmann, E. E. (1996). Functional vision assessment of infants. In B. Rosenthal & R. Cole (Eds.), *Functional assessment of low vision* (pp. 45–62). St. Louis, MO: Mosby.

Hartmann, E. E. (2000). Visual functioning in pediatric populations with low vision. In B. Silverstone, M. Lang, B. Rosenthal, & E. Faye (Eds.), *The lighthouse handbook on vision impairment and vision rehabilitation.* New York: Oxford University Press.

Isenberg, S. J. (Ed.). (1994). *The eye in infancy.* St. Louis, MO: Mosby.

Mayer, D. L., Beiser, A. S., Warner, A. F., Pratt, E. M., Raye, K. N., & Lang, J. M. (1995). Monocular acuity norms for the teller acuity cards between ages one month and four years. *Investigative Ophthalmology & Visual Science, 36*(3), 671–685.

Moore, B. D. (Ed.). (1997). *Eye care for infants and young children.* Boston: Butterworth-Heinemann.

Simons, K. (Ed.). (1993). *Early visual development: Normal and abnormal.* New York: Oxford University Press.

Teller, D. Y. (1997). First glances: The vision of infants. The Friedenwald Lecture. *Investigative Ophthalmology & Visual Science, 38*(11), 2183–2203.

PERCEPTUAL DEVELOPMENT: VISUALLY GUIDED REACHING

Vision plays a central role in the planning and execution of goal-directed actions by informing the actor about the pertinent physical characteristics of the target object. When reaching for a stationary target, for instance, we look at its location and physical properties (size, orientation, shape, etc.) before initiating our movement. If the object is moving, we follow its trajectory visually to anticipate where and when to intercept it. Children and adults who are experienced at reaching and have acquired good control of their arms and posture can easily take advantage of such visual information to tailor their actions effectively to the physical world. But infants who have limited knowledge about the physical world and little control of their body and limbs need to learn how to map their actions to the world. *Visually guided reaching* in infancy has been extensively studied because it represents one of the earliest forms of goal-oriented movement in development. It allowed scientists to address questions about the origins of visuo-manual coordination and assess how young infants use vision to plan and control their actions. This body of work revealed several steps in the development of early eye-hand coordination across the first year of life. This entry on visually guided reaching focuses on newborns' eye-hand coordination and the role of vision in voluntary reaching of newborns.

Eye-Hand Coordination in the Newborn

In the mid-1970s, researchers discovered that neonates can display a rudimentary form of eye-hand coordination. When provided with stable head and trunk support, newborn infants can extend their arm in a forward motion with the hand fully opened. Such full arm and hand extensions have been observed when objects are present in the infant visual field, but also when objects are absent from the infant visual field. These observations raised the question of the intentionality of this neonate response, but careful research from Claes von Hofsten indicated that visual attention influences early arm movements, suggesting that a nascent form of intentionality may already exist from birth. He found that when newborns were fixating the object in front of them, they tended to direct their arm closer to the object than when they were not fixating the object. Furthermore, when the hand came closer to the fixated object, newborns tended to slow down their movement. However, contact with the target almost never occurred. Because movement directedness and contact with the target are poorly controlled in the neonate, researchers called this response "prereaching" to indicate that it is prefunctional and precursory to the voluntary reaching response emerging a few months later.

By seven weeks of age, the rate of prereaching responses declines, and when prereaching occurs, the hand is no longer open, but fisted. This decline in response does not seem to be linked to a loss of interest in the task, as infants increase their rate of looking at the target during this time. It is possible that infant vision and head control, which improve dramatically by two months of age, temporarily inhibit the production of the prereaching response. Soon after, the rate of forward arm responses toward objects visually perceived increases steadily, and the hand begins to reopen gradually during the movement.

The Role of Vision in Voluntary Reaching

Around three to four months of age, infants' intentions to reach for an object in their vicinity have become quite clear. Infants can contact the fixated target with their hand successfully and repeatedly, but they do so awkwardly and quite variably. Some infants thrust their arms forward at high speed and swat at the object, others produce slow seemingly coordinated arm motions and touch the object delicately. Research on infant reaching has aimed to elucidate the perceptual-motor processes driving the emergence and subsequent progression of this voluntary behavior over the first year of life.

Trajectory Formation and the Process Underlying Visuo-Manual Guidance

When infants perform their first voluntary reaching attempts, they need to figure out how to bring their hand to the toy location. In his pioneering work, Jean Piaget thought that infants solved this problem by bringing their hand closer to the perceived object by glancing alternately at the hand and toy while both are in the child's visual field. This view, that infants initially use vision to guide their hand to the target, prevailed for many decades and was supported by a host of studies that used either displacement prisms or kinematic recordings to assess how vision controlled the trajectory of the infant hand during reaching. Kinematic recordings are a way to capture and describe the motion of limbs through computer analysis. The studies that used this method revealed one particularly strong characteristic of infants' early reaches: That the trajectory of the hand to the toy is not straightforward, but discontinuous, with several changes in movement direction before the hand makes contact with the toy. Such discontinuity in the hand path was considered evidence that infants' early reaching movements were visually guided.

This early work led to a view of the development of infants' eye-hand coordination as a two-phase process in which vision was used distinctively in the planning and execution of reaching movements. A first phase, following the onset of voluntary reaching, was called "visually guided" to indicate that during that period infants used vision primarily to guide their hand closer to the aimed target. The second period, beginning around eight months of age, called "visually elicited," emphasized the fact that by that time infants had acquired better arm control and therefore did not need vision anymore to monitor the trajectory of their arm to the target. Instead, infants used vision to identify the physical characteristics of objects in order to shape their movement accordingly. Research performed in the early 1990s contributed significantly to revision of this view.

Rachel Clifton and colleagues followed young infants weekly from 8 to 26 weeks of age and tested them with objects presented in the light and the dark (in the dark, the objects were sounding and/or glowing). They reasoned that if vision is needed for the emergence of reaching and the guidance of the hand, successful reaches in the lighted conditions should appear before successful reaches in the dark condition. The authors found that infants' first successful reaching attempts emerged both in the light and dark conditions at nearly identical developmental times. This finding provided evidence that infants do not need to see their hand to guide it to the target when they are first learning to reach.

The new consensus among researchers is that infants initially rely on proprioception (the inner sense that provides information about joint position and motion) to direct their hand to a perceived target. Consistent with this new view, findings revealed that the early discontinuous hand trajectories typical of infants' early reaches are not due to visual guidance of the hand, but to the infants' difficulty in calibrating the speed of their movement to the task. Esther Thelen and collaborators discovered that the energetic arm thrusts that infants perform in the direction of the target often introduce involuntary, motion-dependent deviations in the path of the hand, which infants correct proprioceptively by contracting their muscles to break the arm motion and bring the hand back on track to where the target is located. As time and practice accumulate, infants learn to regulate the speed of their movement to the task demands and generate progressively more direct hand paths to the target. Thus, from this more recent work, it appears that, at the early stage of reaching development, infants use vision mainly to locate the target in space, not to guide their hand.

Predictive Reaching for Moving Objects

As soon as infants begin reaching for stationary objects at about three to four months of age, they also begin reaching for moving objects. Research has shown that young infants can predict the trajectory of an object. They anticipate object trajectory by moving their arm and hand before the object is within their reaching space. They also intercept the target successfully, however, at that young age, success rate is higher when the object is moving slower than faster.

Success at predictive reaching may find its root before infants begin to reach. Infants can track a moving target with their head and eyes from birth, and by two to three months they are able to produce a smooth eye pursuit of a moving target and even slightly anticipate the position of a target moving back and forth repeatedly. Tracking and reaching for moving targets improves with age, and by six months old infants are capable of extrapolating the path of objects in accord with the principles of inertia, especially if the path is linear.

Perceived Reachability

Soon after reach onset, infants are capable of visually determining when an object is close enough and reachable, versus when an object is out of reach. When objects are presented at different reaching distances, five- and six-month-olds will accurately scale their own projected prehensile space to the object distance. This scaling is also relative to the infants' sitting abilities. When infants can sit alone, they are able to extend their prehensile space further out by leaning forward with their trunk without losing balance. This work reveals that infants' perceived reachability is grounded in an accurate sense of what they can do with their own body.

Reaching for Objects of Different Orientations, Shapes, and Sizes

Reaching for objects of different orientations, shapes, and sizes requires some adaptation of the movement of the hand before contact with the target to secure a good grasp. Such visually guided anticipatory activity of the hand during reaching was found to emerge between 5 and 14 months of age, depending on the objects used in the study.

Infants' earliest attempts to match the shaping of their hand and fingers to the visually perceived physical characteristics of objects before contact have been observed at about five months of age. When infants are presented with rods oriented either horizontally or vertically, they make crude attempts to adjust their hand accordingly, but it is not until seven to nine months of age that infants clearly align their hand and fingers to the orientation of the rods. Other studies that tested infants with objects that were either convex or concave in shape found similar results, but studies that presented rods at a diagonal angle have not observed clear signs of anticipatory hand orientation before the age of 11 months.

Reaching for objects of different sizes requires the differential use of one or two hands, in addition to finger shaping, depending on the size of the object. The emergence of this differential ability has led to inconsistent findings in the literature. Some studies report that infants can scale their reaching and grasping abilities to object sizes as soon as they begin to reach for objects at four months old. Others presented data indicating a more protracted developmental trend, with infants beginning to display such movement scaling only around 9 months or even as late as 13 to 14 months. Part of the discrepancy between findings comes from the type of sensory information that infants are believed to use when adjusting their reaching response. According to some researchers, infants initially rely on haptic information (the sense of touch) to adjust their grip configuration (fingers and hands) to the size of the objects and only later rely on anticipatory visual information alone to configure their arm and hand movement before contacting the target. Other researchers, in contrast, found that both haptic and visual information in the shaping and planning of reaching movements for objects of different sizes emerge later, when infants are freed from pervasive motor constraints influencing the coupling and decoupling of their arms. Some of these motor constraints arise from infants' successive postural reorganizations over the first year of life as they learn to sit, crawl, and walk. For instance, poor sitting and walking abilities at around 6 and 12 months old have been found to alter infants' upper arm movement decoupling. As a result, they respond more bimanually to small and large objects alike.

Despite these discrepancies, the previously mentioned studies as a whole clearly reveal that infants become more proficient at increasing their visuomotor mapping to objects' physical characteristics as they gain better control of their arms and body. Most of the work reported here relied mainly on infants' motor responses to infer their visual and movement-planning abilities. Little is known about how closely infants pay attention to and scan the visual characteristics of objects when preparing to reach. Very recent work is beginning to address this question by pairing eye-tracking technology with motion analysis during reaching. Preliminary results seem to suggest that infants' varying degrees of attention to the target object have direct implications for the manner in which they reach for objects.

Daniela Corbetta

See also Action and Vision; Embodied Perception; Perceptual-Motor Integration; Proprioception; Reaching and Grasping

Further Readings

Clifton, R. K., Muir, D. W., Ashmead, D. H., & Clarkson, M. G. (1993). Is visually guided reaching in early infancy a myth? *Child Development, 64,* 1099–1110.

Corbetta, D., Thelen, E., & Johnson, K. (2000). Motor constraints on the development of perception-action matching in infant reaching. *Infant Behavior and Development, 23,* 351–374.

Piaget, J. (1952). *The origins of intelligence in children.* New York: Norton.

Rochat, P., Goubet, N., & Senders, S. J. (1999). To reach or not to reach? Perception of body effectivities by young infants. *Infant and Child Development, 8,* 129–148.

Thelen, E., Corbetta, D., Kamm, K., Spencer, J., Schneider, K., & Zernicke, R. (1993). The transition to reaching: Mapping intention and intrinsic dynamics. *Child Development, 64,* 1058–1098.

von Hofsten, C. (1982). Eye-hand coordination in the newborn. *Developmental Psychology, 18,* 450–461.

von Hofsten, C., Vishton, P., Spelke, E. S., Feng, Q., & Rosander, K. (1998). Predictive action in infancy: Tracking and reaching for moving objects. *Cognition, 67,* 255–285.

Witherington, D. C. (2005). The development of prospective grasping control between 5 and 7 months: A longitudinal study. *Infancy, 7,* 143–161.

PERCEPTUAL DEVELOPMENT: VISUAL OBJECT PERMANENCE AND IDENTITY

Two of our most basic cognitive capacities, and milestones in early cognitive development, are *object permanence and object identity.* Object permanence refers to the ability to understand that objects continue to exist even when perceptual contact is lost (e.g., recognizing that a favorite rattle still exists even when it is covered with a blanket). Object identity refers to the ability to determine whether an object currently in view is the very same object or a different object than seen before (e.g., deciding whether the rattle currently in view is the very same rattle played with yesterday). Object identity should not be confused with object recognition; these two processes differ in important ways. Object recognition refers to the ability to recognize whether the object currently in view looks the same or different as an object seen previously. Object identity refers to the ability to determine whether two perceptual encounters involve the same object (the same rattle seen twice) or numerically distinct objects (two different rattles). It is possible to recognize an object as looking novel (or familiar) without making inferences about its identity.

Understanding the development of object permanence and object identity, the subjects of this entry, are critical to theories of object knowledge in infancy. These two capacities lay the foundation for more complex thought and behavior. For example, whether an infant perceives a toy as the same toy as played with previously determines how the infant will think about and act on that toy.

Object Permanence

Early developmental psychologists, like Jean Piaget, believed that infants lacked object permanence until the end of the first year. Most of these studies relied on search methods, and infants' failure to search for hidden objects was believed to be indicative of their lack of knowledge that objects continue to exist when out of view. The rationale was that if infants knew that an object continued to exist when placed under a cover or a blanket, they would search for it. However, with the development of more sensitive research methods, including looking time, reaching in the dark, and simplified search tasks, there is now converging evidence from different laboratories using many different methods indicating that infants as young as 2.5 months possess object permanence, even though infants fail to search for hidden objects until about 8 months. Evidence that infants represent the existence of hidden objects has shifted the focus of research away from the question of whether infants possess object permanence to questions about the nature and content of infants' object representations. Current research suggests that infants possess some basic (or core) information about objects, but that this knowledge changes appreciably with time and experience. For example, even very young infants recognize that objects maintain their numerical identity across space and time, but there are age-related changes in the kind of information (e.g., the spatiotemporal or featural properties of the objects) that infants use to track objects.

Object Identity

Object identity was previously defined as the ability to determine whether two perceptual instances involve the same object or two different objects. However, object identity can be conceptualized as two distinct processes. One process, object individuation, is that of determining numerical identity (how many objects are involved) and the other, object identification, is that of identifying each object by its perceptual attributes (e.g., which objects are involved). Identifying objects on the basis of perceptual attributes requires binding specific features to generic entities. The development of each of these processes is described separately.

Object Individuation

Spatiotemporal information is fundamental to the individuation process. By 3.5 months, infants use discontinuities in speed or path of motion to signal the presence of distinct objects. For example, if a object moves behind one edge of an occluder and then immediately reappears at the other edge too quickly to have traveled the width of the occluder, infants use the discontinuity in speed of motion to conclude that two objects were involved in the event. Young infants can also use spatial information to individuate objects, but this develops later than their ability to use featural information. For example, by 4.5 months infants use form features (e.g., shape, size), but it is not until much later that infants use surface features (e.g., color, luminance) as the basis for individuating objects.

There are probably several factors that contribute to infants' earlier sensitivity to form features. One factor has to do with the nature of the developing visual system. Because color vision is initially quite poor, young infants have difficulty getting good information about color. In contrast, infants' sensitivity to areas of high contrast and to motion-related information presents even young infants with many opportunities to gather information about object form. However, visual maturation cannot fully explain the developmental hierarchy favoring form features. Infants are sensitive to surface features long before they use those differences to individuate objects. Another factor has to do with the perceptual nature of form features. Because form features are amodal, they can be experienced in many different modalities and provide redundancy in information across modalities. Finally, form features are deeply embedded in the physical world, are intimately tied to objects (i.e., they specify the physical nature of objects, the space they occupy, their substance, and how they will move and interact with other objects), and are stable over time. In contrast, surface features, such as color and luminance, are not unambiguously linked to objects or relevant to understanding the way in which the physical world operates (e.g., the color of an object does not predict whether it will fit into a container or support another object). In addition, these features are often perceived by infants as unstable across viewing conditions. Because of these factors, infants do not view color information as particularly salient when tracking objects across space and time.

However, this does not mean that infants can never use surface features as the basis for individuating objects. There is evidence that young infants can be primed, through select experiences, to use surface features at an age younger than they do spontaneously. For example, young infants demonstrate greater sensitivity to color differences in an individuation task if they are first shown events in which color predicts the function of an object. That is, making color functionally relevant can increase infants' sensitivity to color information when individuating objects. There is also evidence that increasing the perceptual salience of surface features can facilitate infants' use of those features in an object individuation task. This suggests that infants' object representations are relatively fluid and may depend on the demands of the task and/or the nature of the situation.

Language plays an important role in infants' capacity to individuate objects. For example, infants who are just beginning to learn count nouns (age 9 to 12 months) demonstrate enhanced performance on object individuation tasks when the objects are labeled (regardless of whether the labels are "real" words or nonsense words) than when the objects are associated with other types of sounds, such as "ooh" or "ahhh." Labels provide a simple, yet conceptually meaningful way to tag and track individual entities.

Recent advances in the application of optical imaging technology into the experimental setting have allowed investigators to begin to explore the neural basis for object processing and individuation in the infant. Functional neuroimaging studies using near-infrared spectroscopy have revealed specific areas in the cortex important to the individuation process. For example, the temporal cortex is activated in response to tasks that require analysis of object features, whereas the parietal cortex is activated in response to tasks that require analysis of the spatiotemporal properties of objects. Further research is needed to explore the functional development of neural areas important to the individuation process.

Developmental changes in infants' capacity to individuate objects depends on a number of factors, including maturation of the visual system, experience with objects in the physical world, language, and conceptual development.

Object Identification

Object identification requires the binding of features to individual entities. Although infants as young as 4.5 months use featural information to individuate objects, at least under some conditions, they do not necessary bind those features to objects. That is, infants may individuate-by-feature but not identify-by-feature. For example, if a green ball disappears behind one edge of an occluder and a red box emerges at the other edge, infants may use the featural differences between the ball and the box to infer that two objects are involved in the event and expect to see two objects when the screen is lowered (as indicated by the fact that they show prolonged looking when only one object is revealed). At the same time, young infants may not hold expectations for what those two objects look like (when the screen is lowered they look equally at a display containing a green ball and a red box or one containing two green balls). There is evidence that infants first start to identify-by-feature during the middle or latter part of the first year. This is consistent with the idea that the representations of young infants are more basic, and less detailed, than those of older infants. Finally, there is some evidence that the developmental progression in infants' capacity to bind features to objects is similar to that of object individuation. For example, older infants are more likely to identify objects on the basis of shape differences than color differences. However, this has not been as well researched as infants' changing capacity to individuate objects, so conclusions remain more speculative.

Object Permanence and Identity in Other Perceptual Domains

Although the majority of research investigating object permanence and object identity in infants has been in the visual domain, these processes are not unique to the visual system. Infants represent the presence of objects experienced through other modalities (e.g., tactile, auditory) and individuate objects on the basis of sound differences. In addition, there appears to be a developmental hierarchy in the type of auditory information infants use to individuate objects. Infants are more likely to use property-rich sounds than property-poor sounds in an individuation task. Property-rich sounds are sounds that are intimately tied to the physical properties of an object and the nature of the physical event in which it is engaged (e.g., the sound of a wooden ball hitting a solid surface); and property-poor sounds are sounds that are more contrived and ambiguously linked to objects (e.g., the ringing of a telephone or doorbell). One possible explanation for this finding is that infants have more experience with property-rich than property-poor sounds. Another possible explanation is that property-rich sounds are more intimately tied to objects (e.g., are dependent on the structure and material composition of an object) than property-poor sounds, and hence are a better predictor of an object's identity.

Teresa Wilcox

See also Attention: Object Based; Binding Problem; Object Persistence; Perceptual Development: Color and Contrast; Perceptual Development: Object Perception

Further Readings

Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science, 3*(1), 2–13.

Káldy, Z., & Leslie, A. (2005). A memory span of one? Object identification in 6.5-month-old infants. *Cognition, 97,* 153–177.

Leslie, A. M., Xu, F., Tremoulet, P. & Scholl, B. (1998). Indexing and the object concept: Developing 'what' and 'where' systems. *Trends in Cognitive Sciences, 2,* 10–18.

Spelke, E. S., Kestenbaum, R., Simons, D. J., & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology, 13,* 113–143.

Tremoulet, P. D., Leslie, A. M., & Hall, G. D. (2001). Infant individuation and identification of objects. *Cognitive Development, 15,* 499–522.

Wilcox, T., Bortfeld, H., Armstrong, J., Woods, R., & Boas, D. (in press). Hemodynamic response to featural and spatiotemporal information in the infant brain. *Neuropsychologia.*

Wilcox, T., Bortfeld, H., Woods, R., Wruck, E., & Boas, D. (2008). Hemodynamic response to featural changes in the occipital and inferior temporal cortex in infants: A preliminary methodological exploration. *Developmental Science, 11*(3), 361–370.

Wilcox, T., & Schweinle, A. (2003). Infants' use of speed information to individuate objects in occlusion events. *Infant Behavior and Development, 26,* 253–282.

Wilcox, T., & Woods, R. (2008). Experience primes infants to individuate objects: Illuminating learning mechanisms. In A. Needham & A. Woodward (Eds.), *Learning and the infant mind* (pp. 117–143). New York: Oxford University Press.

Xu, F. (2003). The development of object individuation in infancy. In F. Fagan & H. Hayne (Eds.), *Progress in infancy research* (Vol. 3, pp. 159–192). Mahwah, NJ: Lawrence Erlbaum.

Xu, F. (2007). Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences, 11,* 400–406.

PERCEPTUAL DEVELOPMENTAL DISORDERS

See Vision: Developmental Disorders

PERCEPTUAL EXPERTISE

The study of *perceptual expertise* addresses the acquisition of perceptual skills that generalize across objects in a domain, such as the ability to recognize birds, to match handwriting samples, or to interpret x-rays or weather maps. Although perceptual expertise in many domains is rare (e.g., few people can match fingerprints), most people acquire some kind of expertise, such as the ability to recognize characters of a writing system or to learn new faces. The majority of studies on perceptual expertise have been done in visual domains, although perceptual expertise can be acquired in other modalities, as found in the work of a sommelier or a music critique. Topics discussed in this entry include: different kinds of perceptual expertise, the study of facelike expertise, facelike expertise with nonface objects, and questions for future perceptual expertise research.

The skills studied in perceptual expertise studies can be distinguished from those typically investigated in research on perceptual learning, which reflects improvements that are highly specific to the trained stimulus and the conditions of training. For example, improvements in detecting low-contrast vertical bars in the right visual field may not transfer to performance with horizontal bars in the same location or to vertical bars in the left visual field. In contrast, a bird expert may learn new species of birds on a trip abroad more efficiently than a novice would. However, the mechanisms underlying perceptual learning may be the same as those involved in perceptual expertise. For instance, perceptual learning could reflect early stages of perceptual expertise, but this is still unknown.

The processes studied under the rubric of perceptual expertise also differ from those that underlie expertise of a more abstract nature, although real-world experts likely generally acquire both perceptual and more semantic knowledge about objects. For instance, a car expert may learn to distinguish cars from photographs, but also learn about the history of the car designs. In contrast to knowledge that can be verbalized, perceptual expertise is not accessible explicitly for the expert to report. A doctor may be able to discuss with colleagues the rules by which she makes a given diagnosis, but her ability to discriminate different kinds of rashes effortlessly is not something that she can explain. Indeed, there is some evidence from studies of face expertise that verbalizing can impair expert performance by leading them to use nonoptimal strategies.

Different Kinds of Perceptual Expertise

It is currently unclear how many different kinds of perceptual expertise there are or exactly how they should be classified, but within the same modality (e.g., visual object recognition) we can acquire expertise in very different ways. A brain surgeon may represent the cortex differently than a neuroscientist who compares brain anatomy across species.

The specific task that the expert has learned to solve determines to a large extent the nature of the strategies that are recruited. For instance, face recognition requires the expert to pay attention to subtle differences between shape parts but also the distances between these parts, because all faces share similar parts in the same general configuration (two eyes above a nose and a mouth). This means that representations and mechanisms that are sufficient for everyday object recognition (e.g., recognizing a chair or a car, which can be done by detecting the presence of certain parts in a given configuration) are unlikely to be efficient for individuating faces. Reading, in contrast, requires one to ignore large differences in the shape, size, and color of letters so they can be recognized and combined to form words. Therefore, to understand expertise, we must understand the task that the expert has learned to solve.

Because there are in principle as many different kinds of expertise as there are possible tasks, it is difficult, if not impossible, to make any general statement about the acquisition of expertise in a domain that would generalize across all other domains. This entry focuses on one kind of expertise that has been the subject of many studies: the recognition of objects that share a configuration of parts, or "facelike" expertise.

The Study of Facelike Expertise

Face recognition, specifically the recognition of upright faces, is considered special compared to the recognition of other objects. While many researchers study face recognition in its own right, others use it as a case study for a kind of perceptual expertise in which people become very good at telling apart visually similar objects.

First, there are brain mechanisms specifically engaged by face recognition, including activity in an area of the extrastriate cortex called the fusiform face area (FFA), which responds more to upright than inverted faces. Second, face recognition leads to a host of behavioral effects that distinguish it from object recognition. The recognition of upright faces depends on configural information, with observers easily detecting small changes in the distance between parts when a face is upright, while the same changes can be difficult to detect when the face is upside down. In contrast, inversion does not have the same impact when observers detect feature changes rather than configural changes. The importance of configuration for faces can be seen in the "Thatcher effect" (first demonstrated on an image of Margaret Thatcher). In this demonstration, face parts (e.g., the eyes and the mouth) are turned upside down. This gives the face a grotesque expression, but what is striking is that this transformation is almost undetectable when the entire image is upside down.

Another hallmark of face recognition is "holistic processing," or the tendency to process all features of a face together, rather than independently. This can be demonstrated as an impairment in the ability to identify a single face part accurately when other parts of the face are changed, even when observers are warned of the manipulation. For instance, when asked to decide whether the nose in an image is that of Elvis, observers make more mistakes or are slower if the eyes or mouth in the picture are replaced by those of Mick Jagger. Phenomenologically, it seems that all the parts of the composite face fuse together, taking the appearance of a new face that looks like neither of its "parents." Holistic processing is influenced by the configuration of the parts: It becomes much easier to pay attention to one part and ignore the other parts if the parts of the composite face are misaligned in an abnormal configuration.

Facelike Expertise With Nonface Objects

To psychologists interested in perceptual expertise, the possibility that face perception differs from object perception because of expertise is exciting, given that face perception has been extensively studied. For this reason, a great deal of work with expert observers in perceptual expertise consists of trying to replicate face-specific effects with nonface objects.

Over the last 20 years, many studies have considered whether putatively face-specific effects can be obtained for nonface objects provided expertise. One of the first such demonstrations was by Rhea Diamond and Susan Carey in 1986, in which an inversion effect was found for dog images in dog show judges. At the time, it was proposed that it took about 10 years of training to acquire perceptual expertise. Later work by Isabel Gauthier, Michael Tarr, and colleagues provided evidence that facelike effects could be observed for novel objects after only a few hours of training. This is important not only because it suggests more functional plasticity than first thought, but it also demonstrates the feasibility of studies of perceptual expertise in the laboratory, affording more precise control over training conditions and stimulus properties. Next, this entry briefly reviews some of the facelike effects that have been obtained in object experts.

Functional neuroimaging studies reveal that tasks involving experts with nonface objects (cars, birds, novel objects) can engage face-selective areas in the visual system, in particular the right and left FFA and, in some studies, the occipital face area (OFA). The recruitment of face-selective areas is relatively independent of the task performed during the brain scan, as objects of expertise activate these regions both during identification judgments and tasks that do not require attention to the identity of objects (e.g., location judgments). However, not all functional studies of object experts recruit the FFA. Some of this work may have tested experts with objects that are just outside of their domain of expertise. For instance, the skills of a modern car expert may not be optimized for the perception of antique cars. And it is unlikely that all kinds of expertise recruit the FFA; for instance experience with letters recruits other parts of the visual system. But the fact that some kinds of expertise recruit the FFA suggests that this part of the visual system performs computations that are well suited to the rapid discrimination of visually similar objects. Moreover, if we can understand how we acquire face expertise, we may be able to train better experts in other domains.

Holistic and configural processing effects have also been observed in experts with nonface objects. The large cost caused by inversion in face perception is rarely matched in size in studies with nonface objects. However, the inversion effect can show an effect of expertise, suggesting that with enough expertise, the sensitivity to inversion grows. Indeed, in recent work measuring how many faces and cars can be memorized across a short delay, more upright than inverted faces could be memorized, and a similar effect of inversion was found for cars. Even more important, the magnitude of the inversion effect was directly correlated with a separate measure of car expertise. These results also suggest that experts may encode objects of expertise more efficiently, allowing them to hold more of them in the short term memory. Although such an advantage generalizes to new objects within the domain of expertise, it is very specific to the familiar orientation of the objects in that domain.

The inversion effect is only an indirect measure of holistic and or configural processing. However, more direct measures of configural and holistic processing increase with expertise. For example, judgments about one part of an object can be made independently from the identity or configuration of the other parts in novices, but experts are unable to selectively attend to a part of an object of expertise, as with faces. In the case of fingerprint expertise, configural processing has been demonstrated by showing that performance for a complete fingerprint is more than double what is obtained by doubling the performance for two fingerprints filtered to each contain 50% of the original image. This effect, only obtained with experts, suggests that there is additional information in the relation between parts that are used only by experts.

Questions for Future Research

The study of perceptual expertise is relatively young, and some of the general questions in the field of expert object recognition have been overshadowed by the role of expertise in the debate on whether faces are special. Most studies of facelike expertise with nonface objects have focused on showing that hallmarks of face perception can be found in object experts. But few studies have manipulated training protocols systematically to understand how each aspect of the training experience influences performance. Yet such manipulations will be necessary if we want to understand the differences between novice and expert perception, and how we can train perceptual experts. Finally, the relationship between perceptual expertise and other types of learning, such as perceptual learning and the acquisition of facts about objects in a domain, need to be explored.

Isabel Gauthier

See also Agnosia: Visual; Face Perception; Face Perception: Physiological; Object Perception; Perceptual Development: Face Perception; Perceptual Learning

Further Readings

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences, 10*(4):159–166.

Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research, 45*(4), 431–448.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*(2), 107–117.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191–197.

Gauthier, I., & Tarr, M. J. (1997). Becoming a 'Greeble' expert: Exploring mechanisms for face recognition, *Vision Research, 37,* 1673–1682.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11*(1), 8–15.

Tanaka, J. W., Curran, T., & Sheinberg, D. L. (2005). The training and transfer of real-world perceptual expertise. *Psychological Science, 16*(2), 145–151.

PERCEPTUAL LEARNING

Most people are all familiar with the phenomenon of having their performance improve over time with practice. *Perceptual learning* refers to the changes that take place within sensory and perceptual systems as a result of practice performing a perceptual task. Such changes can take place at both the levels of behavior and physiology, and recent work has demonstrated that adult perceptual systems can exhibit a large degree of mutability as a result of extended stimulation through training. This entry explores the background, psychophysics, physiology, and complexity of perceptual learning.

Background

Until relatively recently, it was generally assumed that perceptual learning was a phenomenon that was restricted to the early stages of human development or was attributable to changes in higher-level cognitive processes. In the case of development, a great deal of neural tuning and reorganization takes place during early childhood, and many experiments have shown that perceptual experience (or lack thereof) during this time can play a large role in permanently shaping the properties of neural mechanisms. After this critical period during perceptual development passes, it was traditionally assumed that neural mechanisms at the earliest stages of information processing were no longer plastic and thus could not be modified through experience with the world. In the case of adult perceptual learning, it was generally assumed that changes in high-level cognitive processes (e.g., shifts in decision criteria, attention, strategy) were responsible for improvements in perceptual performance with practice. However, an explosion of recent work has demonstrated that there is, in fact, a great deal of mutability in adult perceptual systems, and these changes often occur at very low levels of processing.

Many of the tasks that we might think of where practice produces improvements in performance are relatively complex, such as learning how to drive a car or ride a bike. There are certainly perceptual components to the learning that is taking place in such complex tasks. But it turns out that our ability to perform extremely simple perceptual tasks with simple stimuli can also dramatically improve with practice. It also turns out that, with the right kinds of psychophysical and physiological experiments, these simple tasks and stimuli can tell us a great deal about the changes that are occurring within a perceptual system as learning is taking place. As such, the majority of research on perceptual learning has focused on these simple tasks and stimuli.

Psychophysics

Psychophysical techniques are designed to allow one to make inferences about the inner workings of a perceptual system just by observing the responses that the system as a whole makes to carefully constructed stimuli. Psychophysical techniques have been used extensively to try to identify the kinds of processing changes that take place with practice in a wide variety of perceptual tasks. Many of these tasks are used because they are thought to tap into the workings of relatively basic perceptual mechanisms. As an illustrative example, consider the extensive amount of work that has been devoted to exploring the impact of training on performance in a task called *vernier acuity*. Imagine a horizontal line drawn in front of you; now imagine splitting that line into two equal halves and slightly shifting one of the halves up or down relative to the other half. How much would you need to shift the line so that you could just barely detect that it no longer lined up with its neighbor? Under optimal conditions, the amount of displacement that you would need to perform this task is shockingly small—significantly less than the size of a single photoreceptor in your eye. What's even more amazing is that performance in this task can greatly improve with practice.

So what does this kind of result tell us about perceptual learning? On its own, perhaps not than much. If we still wished to adhere to the old dogma that there is no plasticity in the earliest stages of human visual cortical processing after childhood development has taken place, we might try to account for this result by positing various high-level changes in visual information processing. But what if we make only one small, seemingly insignificant change to the task: Instead of using horizontal lines, we make the lines vertical. Surprisingly, if we take people that have been extensively trained with horizontal lines and test them with vertical lines, we typically find that the learning does not transfer across tasks: Performance with the vertical lines is usually only marginally better than performance was when people first started training with the horizontal lines. It is only with practice with the vertical lines that performance improves to the same level that was achieved with the horizontal lines. What's more, additional experiments have shown (a) that there is a similar degree of specificity for the position of training (e.g., training in the left visual field does not transfer to the right visual field) as well as the eye of training (i.e., training in the left eye does not transfer to training in the right eye); (b) that explicit accuracy feedback is not required for learning to take place (although the learning process takes place more gradually over time without feedback); (c) that there are multiple phases to the learning process—an initial fast learning phase and a subsequent slower learning phase; and (d) that the learning effects tend to be relatively long lasting, with performance staying nearly as good for weeks or even months after the initial training has taken place.

The kind of specificity of learning found with training in vernier acuity tasks is very difficult to account for by a generalized, high-level learning process; otherwise, why would the learning not simply transfer across orientations, positions, or eyes? It is far more likely that the learning is occurring at a level of neural processing that is tuned to a narrow range of orientations, a particular region of the visual field, and is dominated by input from one eye. As a result, we can make an educated guess as to the physiological locus of learning in a vernier acuity task: Mechanisms with this kind of precision are found only in the primary visual cortex, the very earliest stages of visual cortical processing.

The previous example is meant to illustrate how a small set of well-chosen psychophysical experiments combined with a simple task and stimulus can tell us a great deal about the kinds of changes that are taking place with practice, as well as allow us to form a reasonable hypothesis about their physiological loci. This same general approach has been applied to a wide array of tasks and stimuli that span the sensory modalities. Each of these applications is designed to uncover the underlying changes that take place with practice within a particular kind of perceptual processing (e.g., visual motion detection, tactile spatial discrimination, auditory frequency discrimination). Although there are notable exceptions, the general trend across these tasks and stimuli has been similar to that found with vernier acuity, namely that there tends to be a high degree of task and stimulus specificity of learning.

Physiology

The results of psychophysical experiments suggest that perceptual learning tends to occur at relatively early stages of neural processing. Accordingly, a complementary stream of physiologically based research carried out mostly on nonhuman animals has been focused on localizing and characterizing the neural changes that take place with perceptual learning. Consistent with the results of psychophysical experiments, physiological experiments carried out across all different sensory modalities have found that performance improvements in simple perceptual tasks tends to modify neural mechanisms at early stages of cortical processing. These experiments have also found there to be a great deal of task specificity in the changes that take place in neural representations with learning, such that neural representations that are altered with learning do not exhibit the same sorts of modified response characteristics when presented with different tasks or stimuli.

But how does learning change neural representations, exactly? Experiments have revealed that there are actually several different ways that neural representations are altered by practice in perceptual tasks. The particular changes that take place appear to reflect the specific kind of code used by the brain to represent percepts in a given task. One such change is an increase in the size of the neural representation. With this kind of change, the number of neurons that respond to a stimulus in a given brain region increases as performance in a behavioral task improves. Such changes have been found for a number of tactile discrimination tasks (e.g., two-point discrimination), where learning can produce marked increases in the amount of somatosensory cortex devoted to encoding a particular region of the body (e.g., a finger). Similar changes have also been found in the auditory cortex for auditory discrimination tasks (e.g., frequency discrimination) and in the motor cortex for motor learning tasks (e.g., reaching and grabbing). This kind of change in the neural representation most likely reflects a computational code that relies on summing across a large number of neural responses in order to increase the statistical reliability of an eventual decision.

A second kind of neural change often seen with practice is a sharpening of neural "tuning functions." A tuning function describes the relative sensitivity of a neuron to variations along a particular stimulus dimension (e.g., orientation, frequency). Neurons situated at early stages of perceptual processing generally respond best to a limited range of stimulus attributes, and learning in some cases can serve to narrow the focus of this range. The result of this kind of change is that neighboring neurons will have tuning functions that have less overlap in their responses to stimuli after learning has taken place. Such changes have been found to take place in the visual, auditory, and motor cortex, and likely reflect a code where each neuron produces a response that is as different as possible along a particular stimulus dimension or dimensions (often called *decorrelation*). In some cases, these kinds of changes are also accompanied by a *reduction* in the size of the neural representation. This shrinkage in representation takes place presumably because the narrowing of tuning functions effectively increases the distance among neurons along the dimension that has been trained, thus reducing the total number of neurons that respond to a given stimulus.

A third kind of code that is used by perceptual systems to represent learned information is a change in the relative timing of responses made by a set of neurons. In particular, several studies involving tactile and auditory learning have found that practice discriminating stimuli that vary in their temporal characteristics can produce an increase in the synchronicity of firing across the ensemble of neurons that normally respond to the stimuli. Increased synchrony of neuronal firing has also be found in olfactory learning tasks in which the stimuli are not temporally varying, indicating the use of temporal coding strategies by perceptual systems is not restricted to temporally varying stimuli.

Complexity

As previously described, the majority of research on perceptual learning has focused on using simple tasks and stimuli, and, for the most part, this has been for methodological reasons. However, some of the very latest research on perceptual learning has also employed more complex tasks and stimuli, such as faces and objects. A general finding from these experiments has been that more complex tasks and stimuli tend to produce greater learning effects and show greater amounts of transfer of learning to other tasks and stimuli. Such results presumably reflect the effects of learning on higher-order representations that are tuned to more abstract stimulus properties than those found at the initial stages of cortical processing. Thus, a challenge for future work on perceptual learning will be to further specify the psychophysical and neural mechanisms that mediate perceptual learning in contexts with a greater degree of stimulus and task complexity.

Jason M. Gold

See also Cortical Organization; Experience-Dependent Plasticity; Perceptual Expertise; Psychophysical Approach; Visual Acuity

Further Readings

Fahle, M., & Poggio, T. (2002). *Perceptual learning.* Cambridge: MIT Press.

Fine, I., & Jacobs, R. A. (2002). Comparing perceptual learning tasks: A review. *Journal of Vision, 2*(2), 190–203.

Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron, 31*(5), 681–697.

PERCEPTUAL-MOTOR INTEGRATION

We gain information about the world through many senses simultaneously, and part of the information we gain is available through more than one sense. Imagine holding a lemon in your hands and gently exploring it. You both see and feel the shape of the lemon as well as its nubbled texture. If you squeeze the lemon, you feel its resistance and see its shape deform. These kinds of information are available both visually and haptically (through touch). Other senses can participate as well. If we scratch the lemon, we sense its pungent odor. If we drop it, we hear its impact with the floor. The perceptual system combines disparate sources of information into a unified picture of the environment around us. This entry is primarily concerned with how visual and haptic information is integrated during active exploration of the environment.

There are three central research problems associated with perceptual-motor integration: (1) *Fusion:* When different senses, such as visual and haptic, provide estimates of the same property of the world, how well do we combine these sources of sensory information to obtain a more reliable estimate of the property? (2) *Exploration:* How does systematic exploration of the environment alter visual and other sensory input, potentially enhancing its value to the organism? (3) *Calibration:* How does the nervous system work out the correspondences between sensory information in different sensory modalities?

The Fusion Problem

This problem of combining information from different sources is referred to as "sensor fusion" in computer science and robotics and as "cue combination" in psychology. First of all, we can approach it by asking what benefit is there in combining cues from different sensory modalities? The theoretical justification for doing so is statistical: Any sensory estimate is inherently uncertain and, by combining multiple estimates from different sensory modalities, the resulting composite estimate can be more reliable (less variable) than any of the individual estimates that go into it. Michael Landy, Laurence Maloney, and colleagues summarize the theory and pointed out that, because sensory estimates from any modality can vary in reliability, the rule for combining them to get the best composite estimate should also change. If we are judging the shape of an object through vision and active touch in a dimly lit room, the nervous system should give more weight to the haptic information. If the same judgment is carried out under bright illumination, we should give more weight to visual information.

Marc Ernst and Marty Banks carried out an elegant experiment to test the claim that the perceptual system integrates visual information and haptic information so as to produce a composite estimate as reliable as possible. They asked a subject to judge the size of a small object by viewing and also grasping it between his index finger and thumb. However, the experiment was carried out in a virtual environment. The subject in the experiment "saw" the object presented binocularly and, when he moved to grasp it, he "felt" the object as well. However, the pressure that he felt in grasping was exerted by robotic arms attached to his index finger and thumb. The visual impression of the objects was equally unreal, a product of computer graphics.

With this setup, Ernst and Banks could systematically reduce or increase the reliability of visual relative to haptic information and test whether the nervous system correctly formed composite estimates of object size that were as reliable as possible, given the information from the two sensory sources. They found that observers did vary the weight given to visual and haptic information and that their performance was indistinguishable from that of a statistically optimal observer.

Ernst and Banks picked experimental conditions where two sensory modalities signaled simultaneous estimates of the same property of the world, the size of a small object, at just one point in time. Part of what makes perceptual motor integration a hard problem is that judging the shape of the lemon both visually and haptically may involve an extended period of exploration with the motor system affecting what the visual system sees. If, for example, we grasp a lemon and roll it around in our hands for a few seconds, holding it for up for visual inspection from different angles, the visual input is affected by changes in viewpoint under motor control.

Active Exploration

Much research in perception concerns single views of objects or other stimuli with the viewpoint under the control of the experimenter. But in everyday environments we typically move around, changing viewpoint systematically, and we may even pick up objects like the previously mentioned lemon, rotating it so that we see it from multiple viewpoints. J. J. Gibson and Ulrich Neisser have both proposed that systematic exploration enhances performance.

Using the motor system to rotate the object so that it is seen from different viewpoints is itself of benefit to visual recognition. Karin Harman, Keith Humphrey, and Melvyn Goodale showed that allowing subjects to actively explore computer-generated objects by rotating them on a computer screen facilitated visual recognition. One group of subjects were allowed to freely rotate and examine objects, a second group were given the same visual input but viewed passively: They saw what subjects in the first group saw but changes in viewpoint were not under their control. The former group had faster visual recognition of previously seen objects when later tested with the objects they had seen and novel objects.

The Calibration Problem

William James characterized the baby's experience of the world as a "booming, buzzing confusion" (the original has "blooming," which may be a typographical error). Given such an initial state, how is it possible to organize sensory input to allow accurate combination of information from visual and haptic modalities, as Ernst and Banks found, or systematic exploration, as found by Harman and colleagues? One clue is given in a classic study by Richard Held and Alan Hein. Kittens were raised from birth in two environments. Part of the time both kittens were free to move, but they were in darkness, without visual stimulation. The remainder of the time, the first kitten was free to move along a circular track, the other was strapped into a suspended carrier that was pulled along by the first kitten. In this way, Held and Hein tried to equate the visual stimulation experienced by both kittens with one major difference. The first kitten experienced visual stimulation that was the result of its own planned actions. The second kitten experienced very similar visual input but passively.

After a few weeks, the kittens were allowed to leave the cage. The first kitten behaved normally, moving as gracefully as a kitten should. The second kitten moved awkwardly, stumbling and bumping into objects. Richard Held, in other studies, found similar results with humans adapting to prism glasses that inverted their visual field. Subjects who moved around under their own control learned to compensate for the prism glasses, whereas subjects who were passively wheeled about in a wheelchair did not compensate as rapidly or to as great an extent.

The experiment by Held and his colleagues demonstrated that one important factor in learning to combine sensor and motor input is active exploration of the environment with visual input.

Held's results are consistent with a conjecture of Hermann von Helmholtz in the 19th century. Von Helmholtz noted that if he moved his eye by pressing on it, the visual field appeared to move, while if he moved his eye in the ordinary way, the visual field appeared stable, unmoving. He theorized that, in the latter case, the eye movement was the result of a specific motor command to move the eye, and he postulated that a copy of this command was used to compensate for retinal movement. By analogy, if you are standing in a moving bus and can see that it is about to slow down, you can more readily keep your balance than if you are caught unaware.

Von Helmholtz's efference copy theory provides insight into those of Held and colleagues. In the case where movement is passive, there is no efference signal to copy, and consequently no possibility of comparing what occurs to what was predicted to occur, and therefore there is little possibility of learning how movement affects visual perception and vice versa.

The efference copy theory now has expanded to a range of adaptive control system models where the desired movement is compared to actual movement and anticipated sensory input is compared to actual sensory input. Any discrepancies trigger corrective movements. Sarah-Jayne Blakemore, Daniel Wolpert, and Chris Frith propose that one consequence of the predictive nature of action is that you cannot tickle yourself. They show that, when actual sensory input is consistent with predicted sensory input, the sensory input is attenuated. Thus, if we try to tickle ourselves, the close match between predicted sensation eliminates any sense of tickling, yet the same stimulation unanticipated would reduce us to tears.

Laurence T. Maloney

See also Action and Vision; Corollary Discharge; Haptics; Multimodal Interactions: Visual–Haptic; Reaching and Grasping

Further Readings

Blakemore, S. J., Wolpert, D., & Frith, C. (2000). Why can't you tickle yourself? *NeuroReport, 11,* 11–16.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429–433.

Gibson, J. J. (1962). Observations on active touch. *Psychological Review, 69,* 477–491.

Harman, K. L., Humphrey, G. K., & Goodale, M. A. (1999). Active manual control of object views facilitates visual recognition. *Current Biology, 9*(22), 1315–1318.

Held, R. (1965). Plasticity in sensory-motor systems. *Scientific American, 213,* 84–94.

Held, R., & Hein, A. (1963). Movement-produced stimulation in the development of visually guided behaviour. *Journal of Comparative & Physiological Psychology, 56,* 872–876.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35,* 389–412.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology.* San Francisco: W. H. Freeman.

PERCEPTUAL ORGANIZATION: AUDITORY

See Auditory Scene Analysis

PERCEPTUAL ORGANIZATION: VISION

Perceptual organization is part of the process by which sensory data collected by our receptors is combined and integrated into structured, organized percepts. The eyes signal information about the light around us, but humans do not consciously experience these separate pieces of information—we don't see countless patches of color from the wavelengths registered by our cones. Instead, we see whole objects and surfaces smoothly integrated into scenes. These wholes are the products of perceptual organization. Sometimes these integrations contain surprises, where the whole pattern may be strikingly different from the sum of its parts: an artist draws a few curves on the canvas and a face emerges; or a set of black dots comprising a newspaper photograph combine into a rich scene; or three black disks with wedged-shaped notches yield the perception of a white triangle (see *Perceptual Segregation,* Figure 2). Such novel wholes are called *gestalts.* This entry reviews what perceptual organization does, how it does it, and its practical implications.

A common view sees visual perception beginning with the registration and transduction of light within the retina into neural signals, a process that takes place without our conscious awareness. This is followed by the organization and interpretation of those neural signals into structured, conscious percepts. These processes are sometimes called *sensation* and *perception,* respectively. There seems to be no clear border between them, however; rather, sensation flows smoothly into perception. The latter explains why we usually see whole scenes rather than swirls of megapixels or innumerable separate pieces of a jigsaw puzzle. Determining how we perceive wholes rather than just local parts has proven to be a challenge.

Perceptual organization is an important part of that second process in vision because it deals with how those elementary pieces, sometimes called basic features, are assembled into organized wholes. A commonly held view is that while the early phase of vision proceeds automatically and in parallel (simultaneously) across the visual field, the integration of the component features requires attention and takes place only within a limited region of space at any one time, within a spotlight of attention. If a red ball is tossed our way, one part of our visual system processes the wavelengths leading to red, another detects the curved edges defining a ball, another processes the motion of the ball toward us, and yet others detect the ball's size, texture, and spin. These components are subsequently combined into final percepts through attention, a process called feature binding.

Gestalts

Perceptual organization specifies a very different way parts combine into integrated configurations and how those parts and wholes can be identified. The structuralists, such as Edward Titchener (1867–1927), held that the combination was additive: When the parts are perceived, their sum defines the whole. In this manner, the perception of a dog equals the sum of the percepts of its parts: four legs, a torso, one tail, a head with eyes, ears, nose, and mouth, and fur of a certain length and color. Gestalt psychologists rejected this notion of additivity, claiming that "the whole is different from the sum of its parts" (often misquoted as "more than the sum"). They noted that wholes often possess configural qualities or patterns—gestalts—and that the relationship among the parts was often more important than the parts themselves (e.g., a melody is defined by the relationship among its notes and can thus be played in any key).

The Gestalt movement was launched in the early 1900s with the study of apparent motion, an illusion arising when two stationary lights are flickered alternately, as in movie marquees or in roadside warning signals. Gestaltist researchers produced demonstrations showing that when parts or elements are combined in certain ways, something novel emerges from their configuration, often something surprising. Similarly, they demonstrated that a given part may assume one appearance in one configural context but another in a different context.

Consider the Gestaltists' apparent motion stimulus. We start with a flashing light. Then we introduce a second light flashing in the opposite rhythm, such that whenever one light is illuminated the other is extinguished. When looking at both lights, we see something different from the sum of what we saw from the two lights individually: We now perceive a single, nonflashing light moving back and forth between the two locations. (Seeing just one light with no flashing illustrates how the whole can be less than the sum of its parts.) Here, motion is an emergent feature, something categorically different and surprising because it is a property of neither light individually. The whole we perceive—the configuration or gestalt—differs from the sum of its parts.

Many essential issues of perceptual organization are apparent in Figure 1(a), which shows an old, weathered photograph that is hard to interpret. We can sense the image well enough, in that we can report the gray level of any spot in the picture; we could even duplicate it with a paintbrush, but we might still fail to recognize the familiar object it depicts.

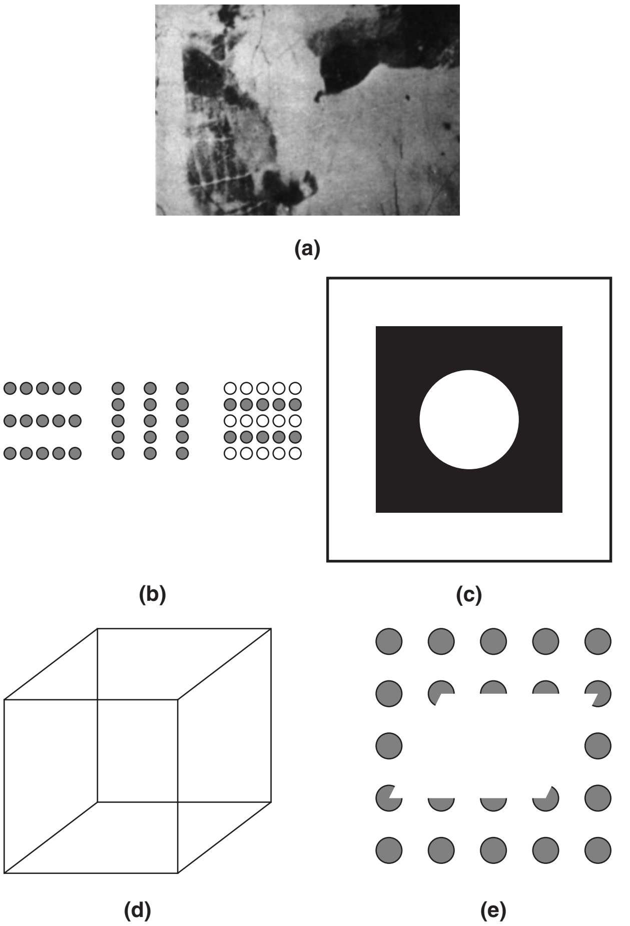


Figure 1

Notes: (a) What's in this image? (b) Grouping by proximity and similarity (c) A circle on top of a square, or a square with a circular hole? (d) Necker Cube (e) Do we see the illusory parallelogram because it is simpler or more probable?

Four Key Phenomena

Perceptual organization focuses on four key phenomena:

1. *Grouping and part-whole relationships:* determining which regions of an image go with which others to form unitary objects.
2. *Figure-ground segregation:* determining which regions represent opaque objects blocking our view of ("occluding") other, more distant objects; and which side of an edge is the figure side and which side belongs to the ground continuing behind.
3. *Perceptual coupling:* determining the appropriate relationship between two linked dimensions in the image. As an object moves away from us, the image it projects to our eye shrinks until it has vanished. If a medium-sized image strikes our retinas, did it come from a large object at a great distance, a small object at a short distance, or an intermediate-sized object at a moderate distance?
4. *Multistability (bistable perception):* some stimuli may be perceived equally correctly in two different ways. Interestingly, our visual system often alternates spontaneously between possible interpretations, abruptly and unrelentingly flipping as though the stimulus were changing.

Let us examine these four phenomena in greater detail, with examples.

Grouping and Part-Whole Relationships

Today's best high-definition televisions have a resolution of about 2,073,600 pixels. The picture comprises a 1920 × 1080 grid, with each cell containing a number indicating how bright the picture is at that spot. (Even better, think of the picture as a stack of three grids or spreadsheets, representing the three color channels). The same is true for a newspaper picture: a simple grid of closely spaced dots varying in size. You may think of the image leaving the eye for the brain similarly—a grid of neurally coded intensity values, ready to be recognized as objects in the cortex. The challenge of grouping is to specify which pixels go with which others to form those objects. Neither your television set nor your digital camera solves this problem; they have no idea what is in the image they are transmitting! Humans, however, perform grouping continuously, accurately, and effortlessly whenever our eyes are open.

To see how to solve this problem, think about solving jigsaw puzzles. Faced with a jumble of pieces, you search for ones with similar colors and for edges continuing smoothly from one piece into another. Research indicates our visual system works similarly, using principles to determine grouping. Some common grouping principles are:

- *Proximity:* the closer together any two elements are in an image, the more likely they belong to the same object (see Figure 1b).
- *Similarity:* the more alike any two elements are (more similar in color, size, orientation, distance, etc.), the more likely they belong to the same object (Figure 1b).
- *Common fate:* the more similarly any two elements change over time (e.g., in their pattern of motion) the more likely they belong together.
- *Good continuation:* the more smoothly one edge or contour blends into another one, the more likely they are parts of a single contour.
- *Closure/convexity:* when connecting contours into objects, curves that can be assembled into closed or convex objects are more likely to belong together than ones that cannot.
- *Common region:* any two elements that are contained within a common region (e.g., encircled by a single contour) are more likely to belong to the same object.
- *Connectedness:* any two elements that are physically connected to one another are more likely to be parts of the same object than two that are not.

These principles operate like rules of thumb or guidelines rather than laws, allowing us to connect picture elements appropriately and to segment (parse) an image correctly. Consider looking out at a sea of faces in a crowd. If we could not group correctly the various facial features that belong to a single person—the eyes, ears, nose, and mouth—we might end up trying to recognize an individual by combining the right eye of one person, the left eye of another, the nose of a third, and the mouth of a fourth. Instead, our visual system may use principles like proximity, similarity, common area, and connectedness to group the features correctly, all without our conscious awareness of these principles. When a long snake slithers through the grass and under a log, we can recognize that snake as a single object because its components are close together, because the coloration across its body is relatively uniform and different from the background, because the contours of its body continue smoothly without sharp angles, and because the whole animal moves together.

Figure-Ground Segregation

Much of the information in an image is conveyed by the edges it contains. Line drawings and cartoons depict scenes using only those edges. Looking around, we see edges everywhere—the edges of our desk and chair, of our arms, hands, and fingers, edges of the room we are in and of the trees outside the window. When our visual system encounters an edge, it has some figuring to do, including whether the edge comes from one object occluding our view of another—as when you put your hand across the sleeve of your shirt. (Other edges arise when two objects merely abut, as with tiles; or when two surfaces come together, as with a corner in a room, or from shading.) With occlusion edges, the visual system tries to determine which side of the edge belongs to the nearer, occluding object (i.e., the figure) and which does not. When your hand covers your sleeve, it is important that the brain correctly assign the edge where they meet to your hand and not to your sleeve, which continues unseen behind your hand and whose own edges have nothing to do with the edges your hand leaves on it.

Imagine you see a circular black patch on an otherwise white wall. Is the round region a black disk that has been taped or painted onto the wall, or is it a hole passing through the wall into the dark? See Figure 1(c) for a similar illustration of the puzzle of figure-ground segregation. If we make the assignment incorrectly, we may fail to recognize objects, as when one looks at a map of Europe and mistakes seas and oceans for land. Escher produced many drawings playing with these phenomena, and the Federal Express® logo famously contains a figure that is hidden until you reverse your figure-ground assignments.

As with grouping, several principles help us make the correct assignments:

- *Convexity:* the convex side of a region is most likely the figure. This helps explain why a dark circle on a wall is more often seen as a painted circle than as a dark hole.
- *Size:* the smaller region is more likely to be the figure. This too helps with the circle.
- *Motion:* the moving region is more likely to be the figure.
- *Symmetry:* the symmetric or "good" region becomes the figure.
- *Distance:* the nearer region becomes the figure.

Perceptual Coupling

The separate components or features of a stimulus are often coupled physically, such as the size of the image an object projects to the retina and its distance from the perceiver. Our thumbnail at arm's length may easily cover the moon, but we are not fooled into thinking the two are the same actual size. These two dimensions are coupled in our perception too, as is well demonstrated in Emmert's law. Suppose we create an afterimage by staring at a bright square and then look at a nearby brick wall. We see the afterimage of that square on the wall, as though it were really there. If we now look at a more distant wall, we again see the square, but now it appears larger, in that it covers more bricks. If we look at our hand, the afterimage will appear to shrink to fit in our hand. Similarly, if we look at a wall that is oblique to our line of sight, the square afterimage becomes trapezoidal, demonstrating the perceptual coupling now between perceived shape and perceived orientation. These demonstrations reveal how the various components and dimensions of a whole stimulus are organized and linked, rather than perceived independently.

Multistability

Many stimuli entering the eye may be interpreted in multiple ways: an ellipse can also be perceived as a circle viewed from an angle, a diamond may be seen as a tilted square, and a triangle may be seen pointing in any of three directions. For some stimuli, rather than sticking with one interpretation, the visual system cycles though different ones. Despite our conscious knowledge that the stimulus is unchanging, our perception of it flip-flops, creating a rare, peculiar, and often amusing experience. The most famous demonstration of multistability comes from the Necker cube (Figure 1d), a simple line drawing of a wire cube whose perceived orientation reverses nearly uncontrollably every few seconds as it is observed. Multistability arises too in motion perception: The red and blue stripes of a barber pole seem to be moving upward as the pole rotates, despite the fact that they are moving horizontally. If the pole is made short and wide, however, the stripes now appear as they should, moving horizontally. If the pole is made just as wide as it is tall, however, perceivers can now see motion in a variety of directions, with our perception shifting incessantly.

All four of these phenomena have implications for how we perceive our visual world. Some lead to illusions; for example, two identical grey surfaces seem to have different shades if we alter our perception of the surfaces' orientation toward a light source. Similarly, when two regions of a stimulus are perceptually grouped, it is difficult for us to attend selectively to just one; and it is easier for us to make judgments of two parts of a single object than of one part each of two objects.

Theoretical Perspectives

Some theoretical notions have been advanced to account for the four phenomena of perceptual organization. One is the claim that gestalts or wholes are primary in perception, with information about parts being secondary. The claim is that we see the forest before the trees, or that we perceive and remember a melody without noticing the particular notes or the key in which it is played. There is good evidence that with some wholes this is what happens (e.g., people are better at discriminating arrows from triangles than they are at discriminating the difference in orientation of a single line segment that changes an arrow into a triangle). People—and animals—are often better processing holistic properties of stimuli (such as closure) than more primitive, localized properties.

Historically, there have been two overarching hypotheses about why our percepts are organized as they are. The first, generated by the Gestaltists, is the Prägnanz principle, which holds that we organize our percepts in the simplest way possible, much as a soap bubble assumes the simplest possible shape—a sphere. Thus, we perceive an ellipse as a circle at an angle because a circle is simpler than an ellipse (having only one parameter, its diameter, compared with an ellipse's two). The second hypothesis is the likelihood principle, espoused by von Helmholtz, which holds that we organize our percepts in the way that most likely reflects the objects in our world. Thus, we perceive trapezoids as rectangles viewed from an angle because rectangles are so much more common in our environment than trapezoids (think about rectangular doors, buildings, signs, containers, etc.). See Figure 1(e) for another example. A possible reconciliation between these hypotheses notes that simple structures are also more common in nature than complex ones.

Practical Implications

Perceptual organization has ramifications for life outside the lab. An example is camouflage, as found in the animal kingdom with protective coloring and in humans within the military. To disguise ourselves to prevent detection by others, we try to group with our background so we don't stand out as figure on ground. We do this by matching our color and texture characteristics to our surroundings and by freezing our motion, lest we group as a unit by common fate. Perceptual organization also figures prominently into the design of instrument panels and computer interfaces. Here, designers try to group controls for similar functions by placing them close together (proximity), by coloring them the same (similarly), and by enclosing them within the same boundary (common region)—as with windows on computer screens. Similarly, we try to draw attention to controls by making their regions stand out as figures, by giving them unique colors, by making them bright and flashing, and by making them move. Perhaps most importantly, perceptual organization helps us perceive the world both accurately and quickly—two virtues that bring great rewards in competitive or dangerous environments.

James R. Pomerantz
and Mary C. Portillo

See also Bistable Perception; Constancy; Context Effects in Perception; Feature Integration Theory; Gestalt Approach; Object Perception; Perceptual Segregation

Further Readings

Garner, W. R. (1974). *The processing of information and structure*. Mahwah, NJ: Lawrence Erlbaum.

Kanizsa, G. (1979). *Organization in vision*. New York: Praeger.

Kimchi, R., Behrmann, M., & Olson, C. (Eds.). (2003). *Perceptual organization in vision: Behavioral and neural perspectives*. Mahwah, NJ: Lawrence Erlbaum.

Köhler, W. (1947). *Gestalt psychology*. New York: Liveright. (Original work published 1929)

Kubovy, M., & Pomerantz, J. R. (Eds.). (1981). *Perceptual organization*. Hillsdale, NJ: Lawrence Erlbaum.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge: Bradford Books/MIT Press.

PERCEPTUAL REPRESENTATION (PHILOSOPHY)

Philosophers and psychologists talk about perceptual states as *representations*. One of the most important questions in the philosophy of perception is whether and in what sense could perceptual states be considered representations. A further crucial question is in what way perceptual representations are different from nonperceptual mental representations, such as beliefs or thoughts. These questions about representations will be discussed in this entry.

Representations

There are various kinds of representation. Leonardo da Vinci's *Mona Lisa* represents a woman: It is a pictorial representation. If you are thinking about Paris, your thought represents Paris: It is a mental representation. Representations refer to things that may be far away (like Paris) or that may not even exist (or, does not exist any more, like the sitter of *Mona Lisa*). They refer to something, they are about something, and what they are about is the object of the representation. The same object could be represented in different ways. The represented object as represented in a representation is called the *content* of the representation.

There are also objects that could be said to represent something without being representations per se. The number of tree rings may be said to represent the age of the tree, for example. The number of tree rings is indeed a reliable indicator of the age of the tree, but it does not literally represent anything. It is a necessary feature of representations that they can misrepresent. Leonardo could have misrepresented Mona Lisa. Your thought about Paris may misrepresent Paris (if you have not been there and you confuse it with Rome). The number of tree rings, in contrast, cannot misrepresent the age of the tree. If there are 20 rings and the tree is only 10 years old, then the number of tree rings does not represent the age of the tree, nor is it a reliable indicator thereof. It does not *misrepresent* the tree's age. It just fails to represent it.

Some of our mental states are representational. Most of our emotions are about something: We are afraid *of* a lion, for example. The same goes for beliefs, desires, and imaginings. It seems natural, then, to suppose that perceptual states are also representations: When you see a cat, your perceptual state is about this cat; it refers to this cat. The content of your perceptual representation is the cat.

Talking about perceptual representations has some important explanatory advantages. Two of the most important philosophical questions about perception are: (1) What is the difference between perception and sensation? (2) What is the difference between perception and belief? According to the standard picture of perceptual processing, mere sensation, that is, the stimulation of our sensory organs, at some point in the processing gives rise to perception, and perception then (sometimes) gives rise to beliefs. In order to know what perception is, we need to have a good way of delineating it from sensation on one hand and beliefs on the other.

If we think of perceptual states as representations, then there is a (relatively) simple way of drawing these lines. Perception is representational, but sensation is not. The stimulation of our sensory organs (of the retina, for example) does not have content: It does not represent anything (it may be a reliable indicator though, like the tree rings). But our perceptual states represent the perceived object as having certain properties. So a clear division line could be drawn between sensation on one hand, which does not have content, and perception and belief on the other, which do. And as perception and belief supposedly represent the world differently, this representational difference between the two kinds of mental state may be used to draw the line between perception and belief.

Perception Versus Belief

But what is this representational difference between our perception and our beliefs? We know how thoughts or beliefs represent the world: If you think that Paris is the capital of France, then you represent something, Paris, in a certain way, as being the capital of France. The question is whether perceptual states represent the world differently. There are a number of proposals about the differences between the way beliefs and perceptual states represent the world. One popular suggestion is that while perception represents the world in a very fine-grained manner, beliefs do so in a rather course-grained fashion: An image is worth thousand words, as they say. This suggestion was criticized, as the content of beliefs can be very fine grained and the content of some of our perceptual states is very course grained indeed.

Another interesting potential difference concerns the indexicality of these representations. Some of our beliefs have *indexical content,* which means that the correctness of the belief depends on the context of the tokening of this belief. Your belief that today is Sunday or that you are now in Vancouver have indexical content, as the correctness of these beliefs depends on when (and where) you have these beliefs. Some other beliefs have nonindexical contents: The belief that Paris is the capital of France does not have indexical content; the correctness of this belief does not depend on when or where you have this belief or even who has this belief. Thus, *some but not all* beliefs have indexical content. But one could argue that the content of perceptual states is always indexical. If you see a cat, you always see it as being in front of you or on your left or on your right: as being localized in your egocentric space. If Bill is sitting across the room from you, he may also see the same cat, but if you see the cat as being on your right, he'll see it as being on his left. Thus, although you and Bill see the same cat, the content of his perceptual state is different from that of yours. The content of one's perceptual state, the argument goes, is *always* indexed to the person who has this perceptual state.

A third potential difference between perceptual content and belief content is that while the content of one's beliefs is conceptual, that is, it could not have the content it has if one didn't master certain concepts, the same may not be true of perceptual states. You could not have a belief that Paris is the capital of France if you did not master the concept of capital. But, arguably, you could perceive a cat without mastering any concept at all, including that of cat. You may not see the cat as a cat (which may require some conceptual apparatus), but you see the cat nonetheless.

Fourth, the content of your belief that Paris is the capital of France is sensitive to the content of your other beliefs. In fact, you would not be able to have this belief unless you had some other beliefs, such as the one that Paris is a city. The same is not true of perceptual states. Our perceptual states can be very insensitive to our beliefs. We know that the two lines in the Müller-Lyer illusion are of the same length, but we can't help perceiving them as having different length. (See Cultural Effects on Visual Perception, Figure 1.)

Finally, a crucial difference between the content of beliefs and of perceptual states is the nature of represented properties. Beliefs can represent their objects as having pretty much any property. Perceptual states, in contrast, represent their object as having a limited set of properties: as having a certain shape, size, color, and spatial location. The list may be extended, but it is unlikely to encompass all properties. You do not perceptually represent the object in front of you as a laptop made in 2006 in Seattle. The question is then which properties are represented in perception and which ones are not.

One way of bringing out the relevance of this question is this. Suppose that you are looking at the duck-rabbit figure: an ambiguous figure that can be seen as a drawing of a duck or of a rabbit, but not both (see Visual Imagery, Figure 1c). Suppose that you experience a Gestalt switch (you saw the figure as a rabbit before, but you see it as a duck now). Is this change a perceptual change? Does the content of your perceptual state change? Or is the content of your perceptual state the same, and it is the belief you attach to this perceptual state that changes? If we opt for the former position, then we have to say that the property of being a rabbit is represented perceptually. Another important candidate for a perceptually represented property is the property of affording a certain action. It has been argued that we perceive objects as affording a certain action: We do not just infer that they do, we literally perceive this property.

Perceptual Representation or Perceptual Relation?

Our beliefs can misrepresent and so can our perceptual states. You may hallucinate that there is a cup of coffee on the desk. In this case, you have a perceptual state that misrepresents. It represents a cup of coffee in front of you but in fact there is no cup in front of you. Hallucinations and illusions on this view are considered to be perceptions that misrepresent their objects.

Although considering perceptual states to be representations may be a natural way of describing our perceptual system, and this assumption dominated both the philosophical and the psychological research on perception, some have recently questioned this entire framework. The proposal is that perceptual states are not representations: Perception is a genuine relation between the perceiver and the perceived object and not between the agent and some abstract entity called "perceptual content." This recent anti-representationalist view of perception is often called the "relational view of perception."

One of the arguments in favor of this relational view is that if we assume that perception is representational, then we lose the intuitively plausible assumption that the object of perception is always a specific token object. According to the relationalist, if you see Liz and then you see Liz's twin sister who looks exactly alike and wears the same outfit, your perceptual state is still different: Your first perceptual state is of Liz and the second is of her twin sister. You may not be able to distinguish these two perceptual states, but they are nonetheless very different. Most versions of representationalism deny this and say that the two perceptual states are the same: They have the same content.

According to the relationalist view, there is no such thing as perceptual misrepresentation, as there is no such thing as perceptual representation. Our perception is always veridical. But then, what happens when we are hallucinating? The relationalist answer is that hallucinating Liz and perceiving Liz may appear very similar, but in fact they are very different: They belong to different "fundamental psychological kinds"—one is constituted by a genuine relation to Liz, the other is not. These two mental states may be subjectively indistinguishable, but this only means that there are limits to our subjective self-knowledge. Hence, it follows from the relational view that perception and hallucination have little in common—this consequence is often labeled "disjunctivism."

Bence Nanay

See also Content of Perceptual Experience; Intentionality and Perception; Naïve Realism; Philosophical Approaches

Further Readings

Campbell, J. (2002). *Reference and consciousness.* Oxford, UK: Oxford University Press.

Dennett, D. C. (1996). Seeing is believing: Or is it? In K. Akins (Ed.), *Perception* (pp. 158–172). Oxford, UK: Oxford University Press.

Dretske, F. (1969). *Seeing and knowing.* London: Routledge.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Martin, M. G. F. (2004). The limits of self-awareness. *Philosophical Studies, 120,* 37–89.

Peacocke, C. (1989). Perceptual Content. In J. Almong, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 297–329). Oxford, UK: Oxford University Press.

Siegel, S. (2006). Which properties are represented in perception? In T. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 481–503). Oxford, UK: Oxford University Press.

Travis, C. (2004). The silence of the senses. *Mind, 113,* 57–94.

PERCEPTUAL SEGREGATION

Most of us take for granted our ability to see the world around us. Rarely do we take time to consider the remarkable accomplishments our visual system must achieve in order to deliver a highly organized and meaningful representation of the world for the purposes of navigating, manipulating, and understanding our environment. For instance, our visual system seems to effortlessly, and instantly, present us with a beautifully structured world that is composed of distinct objects located in three-dimensional (3-D) space from what is actually a sparse sampling of light projected onto the two-dimensional (2-D) surface on the back of our eyes. Appearances aside, however, visual experience is a highly constructive phenomenon that involves dynamic interaction between the visual stimulus and brain structures on a moment-to-moment basis. The fact that we are able to see a single tree in a forest, even though it shares a vast number of similarities with its neighbors, or that we experience a floating red balloon as a single entity rather than as a collection of separate red, round, and moving features, underscores the complexities that must be overcome.

Perceptual segregation reflects the cognitive ability to separate features belonging to the same real-world entity, the grouping of these features into mutually exclusive areas, and their separation from other incompatible features in a manner that promotes a biologically useful representation of the scene. These processes are usually considered to proceed automatically, preceding attentional allocation, and represent the first stages in a hierarchy in which the perceptual system defines parts and objects, prior to imparting meaning to a scene. Although similarities exist between our sensory mechanisms and other physical systems, such as the analogous optical properties of a camera and the eye, the similarities end there, underscoring the profound and substantial transformations that occur in our sensory systems. This entry reviews contemporary theories regarding the role of segregation in visual perceptual organization and the brain mechanisms that are believed to support these abilities. Analogous auditory segmentation processes are discussed briefly at the end and in separate entries on auditory scene analysis and speech perception.

Psychological Underpinnings

Rooted in classic theories from the Gestalt school, perception is believed to arise from emergent processes that result in a whole exceeding the sum of its parts. In Gestalt terms, perceptual segregation equates to the separation of a *figure* from its *background.* This form of parsing information is governed by "rules" (such as whether there is a surround, as well as the size, orientation, and contrast of elements in a scene) that, all else being equal, result in specific figure and background assignments. For example, a surrounded, smaller, higher-contrast, symmetric, or textured region will be seen as the figure, whereas larger, low-contrast, asymmetric, or untextured regions will be seen as the background.

Under this view, *figure-ground* segregation imparts certain categorical interpretations on the scene that shape how perception is organized. These rules of organization describe the figure as lying in front of a background with the separating contour belonging to the figure rather than the background, concepts known as depth ordering and boundary ownership, respectively. More globally, these principles form a relative structure that governs how elements are grouped to produce a holistic perception of the scene (known as the Gestalt laws of grouping).

Therefore, scene segmentation in its simplest form is based on the mutually exclusive separation of areas of visual space that differ on only a single dimension. Although basic stimulus variables such as luminance, spectral distribution (i.e., color), or motion clearly support segmentation processes, texture-based segmentation has possibly received the greatest attention due to the pervasive nature of texture in our visual environments. Beyond this there exists a close correspondence between texture cues (e.g., orientation or spatial frequency) and the basic selective properties of neurons, thereby providing a strong empirical link between perception and the brain mechanisms that support these abilities.

Texture segregation, or the lack thereof, can be easily observed in Figure 1. Here, a tripartite array of randomly oriented X's, L's, and T's are arranged in three successive columns. As is clearly visible, the X's and L's form a highly salient texture boundary, whereas no clear division exists between the L's and T's. Therefore, while these three elements are each easily discriminable on an individual basis, they do not necessarily form distinct texture regions. Segmentation and grouping processes such as these are strongly driven by simple differences in the local spatial relationships among texture elements, and isolating what basic features induce this from texture segregation has been a major focus of research in the study of texture perception.

Boundaries, Regions, and Surfaces

Segmentation processes are a ubiquitous property of vision, however, the means by which segmentation informs the spatial structure of a scene is less clear.

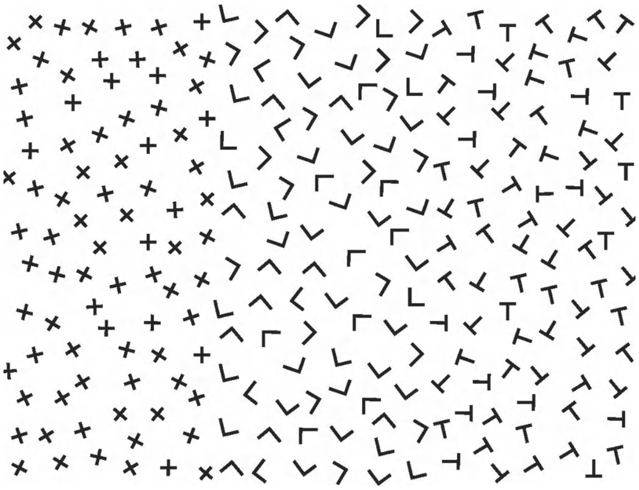


Figure 1 Texture Segregation

Notes: Regions of randomly oriented X's, L's, and T's. Notice that one easily perceives the boundary formed by the X's and L's, whereas the boundary between the L's and T's can only be discerned by effortful item-by-item scrutiny.

The process of dividing a scene into areas based on the uniformity of image-based properties can be achieved either through the detection and separation of differences between image features that are in turn used to divide the scene into distinct areas (so-called boundary-based approach) or through the detection and grouping of similar features (so-called region-based approach).

Under boundary-based approaches, gradients forming a set of interconnected edges are initially detected, which in turn create a contour between the regions. This contour boundary encloses the figure surface and separates it from the background. Region-based approaches, on the other hand, result from processes in which uniform distributions of image-based features are grouped based on their "sameness." Edges are then defined implicitly by the boundaries between these regions. Though both methods provide plausible means by which to accomplish segregation, contemporary theory based on behavioral and neural experimentation as well as computational modeling suggests that the boundary formation processes likely precede region-based processes.

The constructive nature of boundary, region, and surface processes is made more evident when one considers that the visual image projected onto the retina in fact consists of a very large number of localized points in space that possess no explicit representation of which points belong together. Therefore, an important transformation must occur between the image projected in the eye and the perceptual representation of an object in the world. While in fact both the boundary- and region-based approaches previously described rely heavily on the interpolation of spatial features in an image, possibly the best illustration of this transformation comes from illusory contours, as depicted in Figure 2. In this classic demonstration, first put forth by Gaetano Kanizsa, a representation of a triangular surface arises from spatially limited information provided by the three incomplete circles.

The subjective- or illusory-contours that bound the occluding triangle are formed by the perceptual "filling in" of information between local parts. Although the regions inside and outside the triangle have exactly the same physical luminance, the resulting appearance is that of an occluding triangle that is somewhat brighter than its surroundings. Because information propagates over space to create a global perception from the layout of local parts, this example serves to illustrate the constructive properties of vision.

Attention and Perceptual Segregation

As noted, perceptual segregation is viewed as a fundamental problem that the visual system must solve quite early and thus far has primarily been described as proceeding independent of attentional allocation. This view of segmentation as a "preattentive" process is largely supported by findings that grouping and segregation occur without conscious awareness when elements must be distinguished on the basis of relatively simple cues, such as orientation, color, motion, or colinearity. This type of preattentive processing in segregation has also been echoed in research investigating visual search, where a target pattern located in a field of distracter patterns will easily be detected (i.e., will "pop out") if the target and distracter differ at a basic-feature level, such as searching for a red balloon in a sky filled with blue balloons. However, this strictly automatic view of segregation has been called into question by studies showing large decrements in performance on segregation and search tasks when observers must also perform a secondary task that requires effort. In addition, some textures, such as the L's and T's, as shown in Figure 1, can only be discerned slowly by an effortful item-by-item (serial) scrutiny, undermining the notion that segregation is strictly an immediate and effortless process. Therefore, awareness and attention appear to play a modulatory, but not defining, role in perceptual segregation and figure-ground assignment.

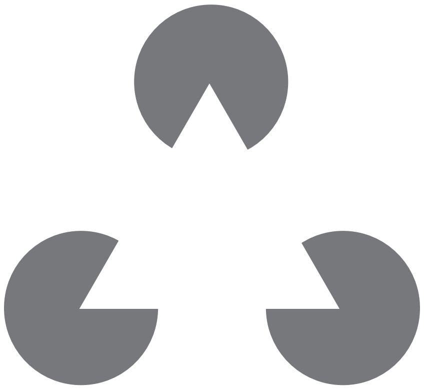


Figure 2 Kanizsa Triangle

Notes: An occluding triangle appears in front of three circles. The triangle appears solid and brighter than the background due to perceptual processes that create subjective contours.

Neural Basis

The neurophysiological correlates of perceptual segregation have received considerable attention over the past two decades, fueled in no small part by the proliferation of powerful analytic methods available to researchers. Such research has focused primarily on how representations of the visual scene arise in the nervous system, given the underlying physiological mechanisms, and the modulatory influence of cognitive processes, such as attention, expectation, and memory.

Theories of the neural implementation of visual perceptual segregation postulate two subsequent stages of processing. In the first stage, basic visual features, such as luminance, spectral distribution, and orientation, are analyzed in parallel in a manner that is preattentive, fast, automatic, and of unlimited capacity. In the second stage, further processing of features, such as boundary determination, surface extraction, and object identification, takes place in a manner that is serial, slow, of limited capacity, and requires attention.

The early stages of visual processing, carried out in the retina, thalamus, and primary visual cortex, are characterized by relatively small receptive fields, and therefore only analyze a small portion of the retinal image. These neurons are selective for the processing of elementary features, and therefore must be combined at later stages (likely through the temporal synchrony in the firing rates of individual neurons) to contribute to perception of the whole. At successively higher stages of processing, more complex features are extracted, in part giving rise to the notion that the visual system has a hierarchical organization.

Different levels of the visual hierarchy do not only receive connections from lower-level areas in the form of feedforward input, but also receive input as feedback from higher-level areas, and even as lateral connections from the same area. This form of recursive interconnectivity between lower- and higher-levels is believed to support the transformations that enable scene segmentation. For example, a neuron low in the processing hierarchy responds more strongly to an oriented line when the line forms part of a figure boundary than when the identical line appears in a homogenous background, even though the receptive field of the neuron is much smaller than the figure. This form of "contextual modulation," in which the responses of neurons in early visual areas are modified by stimuli outside the receptive field, has been supported by lesion studies, showing that damage to higher areas largely eliminates this form of recursive processing and eliminates basic aspects of perceptual segregation.

Perceptual Segregation Beyond Vision

Similar to the problems that must be overcome in translating the 3-D visual world from the 2-D image on the retina, an analogous acoustic problem also exists that must be solved for our perceptual system to generate a useful auditory space. For example, consider the acoustic signal that you may hear when speaking with a friend, while in a room full of other people all simultaneously engaged in conversation. The sound stimuli for all speakers are mixed prior to entering your ear, but you somehow are able to separate out what your friend is saying from all of the other stimuli. This requires processes of segregation and grouping that must overcome the conflation of the mechanical force that causes sound in the world, the differing resonant properties of objects, and the reflection of sound waves that all give rise to a complex stimulus that enters the ear. Beyond this, basic segmentation processes within each sense in turn combine and interact with the output of similar processes in the other senses to provide a rich and detailed multisensory perception of the external world.

L. Gregory Appelbaum

See also Attention: Effect on Perception; Auditory Scene Analysis; Gestalt Approach; Perceptual Organization: Vision; Texture Perception: Visual

Further Readings

Beck, J. (1982). Texture segmentation. In J. Beck (Ed.), *Organization and representation in perception* (pp. 285–317). Hillsdale, NJ: Erlbaum.

Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature, 303,* 696–698.

Kanizsa, G. (1979). *Organization in vision: Essays in Gestalt perception.* New York: Praeger.

Lamme, V. A. F., Rodriguez-Rodriguez, V., & Spekreijse, H. (1999). Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cerebral Cortex, 9*(4), 406–413.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology.* Cambridge: MIT Press.

Rolfsema, P. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience, 29,* 203–227.

PHANTOM LIMB

Healthy, wide awake, and sober, we rarely ever see things, hear voices, or smell odors without a corresponding stimulus. Quite different is the situation after the loss of a limb. Amputees almost invariably continue to feel their lost body parts, sometimes with compelling vividness, in some cases paired with the sensation of mild to excruciating pain. The sensation of a limb that is not physically present is referred to as *phantom limb*. Phantom limbs can teach us very basic lessons about the projective nature of perception. They illustrate the fact that the perception of our own body differs in important ways from the perception of any other object in the outside world. Current research interests in postamputation phantom limbs focus on the functional and structural reorganization of the central nervous system after sudden changes in its communication with the periphery. Phantom limbs after amputation have thus radically changed scientists' views about the adult brain's plasticity: The observation of fast and significant adaptations of cortical areas in response to the loss of a limb has disproved the longstanding notion that maturation would necessarily imply a growing degree of hard wiring of neuronal connections.

Yet, perhaps the most fundamental challenges to theories of body perception come from phantom limbs experienced out of the context of amputation. Phantom sensations of limbs that never physically developed stimulate discussions about innate components of bodily experience. Conversely, patients with lesions to the spinal cord or brain may develop phantom limbs without a physical loss of the corresponding limb, and which are felt spatially separated from their flesh-and-blood counterparts.

Finally, phantom limbs have always inspired philosophical thoughts about the nature of the self and its relation to the body. What happens if we amputate, in a thought experiment, first a finger, then a hand, an arm, and gradually remove even more of the periphery? This question leads us to the concept of a "phantom body" and to currently unresolved issues of how the brain generates, not only out-of-limb sensations, but full-blown out-of-body experiences. This entry discusses postamputation phantom limb phenomena, painful phantom limb experiences and cortical reorganization, and other areas beyond amputation and phantom limbs.

Postamputation Phantom Limbs

Reports about phantom limbs predate the times when successful surgical techniques of amputation were available. Since at least the 10th century, soldiers have described that they would still feel the leg they had lost in combat. Such accounts were taken as evidence for a divine resurrection of the limb, in analogy to the purported resurrection of the entire body after death. In the medical literature, the first detailed description of the phantom limb phenomenon appeared in the 16th century. It is accredited to Ambroise Paré (1510–1590), a French barber, whose skills in using knives advanced him to the country's leading military surgeon, not only admired for the success of his amputations, but also for the development of artificial limbs. As Paré's reports remained largely descriptive, they did not stimulate much dispute among professionals, and the topic vanished from the medical literature for more than three centuries. It resurfaced in the late 19th century in the work of the American neurologist Silas Weir Mitchell (1829–1914). Mitchell, also a prolific novelist, was confronted with the phenomenal persistence of lost limbs in the course of the American Civil War. It is to him that we owe the term *phantom limb*.

During Mitchell's lifetime and the first half of the 20th century, one prominent theoretical question was whether the origin of phantom limb sensations, especially those experienced as painful, had to be sought in the amputation stump ("peripheral theories") or rather at the level of spinal cord and/or brain ("central theories"). Propagators of peripheral theories pointed out that manipulations of the stump (e.g., massage, thermal and electrical stimulation, local anesthesia) would sometimes lead to a reduction of phantom limb pain. In fact, when a limb is severed, radical changes occur at the amputation site. These comprise sprouting of axons and the formation of nodules (neuromas) by the cut nerve fibers. Abnormal firing patterns are a consequence, and these were considered as causative factors in the generation of phantom limb pain. However, distorted impulses from the stump will naturally also pass into the spinal column before ascending through higher-order relay stations, such as the brain stem and thalamus, to ultimately reach the somatosensory cortex. That is why, from today's perspective, the question of an exclusively peripheral or central determination of the phantom limb percept is moot. Center and periphery are interconnected, that is, any changes at the level of the stump will necessarily be accompanied by changes in the brain. The focus of current research on phantom limb phenomena is almost entirely directed to the cerebral cortex—arguably as a consequence of the powerful technologies available to directly monitor the brain's reaction to gross changes in the periphery of our body.

Painful Phantom Limbs and Cortical Reorganization

Whereas probably every amputee is familiar with nonpainful phantom limb sensations, estimates of the incidence of specifically *painful* postamputation phantom limbs vary considerably. Although more that 50% of all amputees have experienced a painful phantom limb at some time after the loss of a limb, severe and chronic phantom limb pain is relatively rare. As a rule, phantom limb pain is more prominent in amputees with preamputation pain (related to disease or injury) and with painful sensations in the amputation stump. More than 50 different treatment methods have been described in the literature, ranging from stump manipulations to a broad range of pharmacological interventions and finally to invasive techniques involving thalamic stimulation or surgical removal of cortical tissue. Not one single method has proven effective for all amputees, but almost every method has brought relief in single cases.

The impetus to investigate changes in the cortical representation of the body after amputation came from observations in animals. During the last quarter of the 20th century, evidence was accumulating that monkeys who had lost a finger showed a reorganization of the area of the cortex previously responsible for the processing of tactile input from that finger. Specifically, this area would now process tactile information from the adjacent fingers, a phenomenon called *cortical remapping*. After deafferentation of a whole arm (i.e., interruption of afferent signals from arm to brain), the area of the cortex that would normally respond to stimulation of the hand was found to be responsive to sensory input from the face. Comparably massive reorganization in the somatosensory cortical map could soon be demonstrated in the adult human brain. Noninvasive techniques, such as magnetic source imaging and functional magnetic resonance tomography, allow the monitoring of activity in areas of the cortex known to process sensory information from circumscribed parts of the periphery. Thus, in arm amputees, the areas coding for inputs from the hand and face (adjacent to one another on the brain's surface) were found to have merged, the former area of hand representation having shifted toward that responsive for facial touch. The more severe amputees' rating of phantom hand pain, the greater these functional cortical shifts were in centimeters. This finding indicated that any factor preventing the cortex from remapping would probably also prevent the occurrence of pain. One of these factors is the use of a functional prosthesis. Contrary to solely cosmetic prostheses, functional models allow the amputee to move parts of the artificial hand, either mechanically or by myoelectric impulses, that is, by electrical signals that make a muscle contract upon the willed intention to move. Thus, the functional properties of the lost hand (grasping, picking, pinching) are not entirely lost, and the hand motor cortex is not completely deprived of its ancestral tasks. In fact, it was repeatedly demonstrated that the more frequently a prosthesis is used, the smaller the representational shifts in the primary motor cortex and the less phantom limb pain was felt.

Another way to fool the brain in "believing" that a lost hand is resurrected is achieved with the aid of a mirror, which is placed vertically between the phantom hand and the remaining hand. Movements of the hand are observed in the mirror and perceived as the apparent "resurrection" of the other, lost hand (note that the mirror image of a left hand is a right hand and vice versa). This mirror therapy can alleviate cramping pain in a phantom hand, whose fingers appear to be clenched against the palm in a frozen position. If the amputee mimics this posture with the fingers of the existing hand, which is steadily observed in the mirror as it is slowly opened, the merging of seen and felt images of the absent hand may restore motility in the phantom fingers and thus abolish the painful sensations. Although repeated exercises with a mirror box can reduce an individual's phantom limb pain, the major impact of mirror therapy consists in the rather transient relief it provides during acute bursts of cramping pain.

Beyond Amputation and Beyond Phantom Limbs

What about people born without an arm or a leg—would they experience phantom sensations of their missing body part? Not according to theories of phantom limb as a process of remembering, that is, based on the brain's memory of previous tactile and motor impressions about a limb. However, cases of differentiated and vivid phantom limbs in people with congenital absence of a limb have repeatedly been described since the 19th century.

Currently, it is debatable whether these observations are evidence for innate components of body schema or whether the brain's "mirror system" may induce feeling a limb by the habitual visual observation of others, who move their corresponding body part. The human mirror system integrates action observation, motor imagery, and motor execution and is presumably important for imitation learning and empathic responding.

The previous presence of a limb is not essential for the occurrence of a phantom limb, and the physical absence of body parts is also not a necessary precondition. For example, people who have experienced a complete severance of the spinal cord have lost any sensory awareness of their body below the injury and are also unable to move their legs (paraplegia). These people may experience phantom legs that are often frozen in a position the legs had taken during injury and typically vanish during visual contemplation of the paralyzed limbs. Vision also interferes with experimentally induced phantom limbs in healthy volunteers; here, illusions of a dissociation between physical and felt (phantom) limb positions can be elicited by anesthesia, tendon vibration, or other manipulations. As soon as the subject watches the real limb, the phantom limb illusion breaks down.

Whereas insight into the illusory nature of a phantom limb percept is always preserved in healthy subjects, amputees, and paraplegics, reality monitoring may be hampered in patients with phantom limbs after brain lesions. Such *supernumerary phantom limbs* are felt in addition to a physically present arm or leg, although, as a rule, the latter is compromised to a variable extent by sensory and/or motor impairment. Specifically after damage involving the parietal cortex, delusional elaborations of a paralysis on one side of the body (hemiplegia) are relatively common. Patients would insist, for instance, that their paralyzed arm belonged to somebody else. In some cases, a personification of the entire hemiplegic side is observed, and a dispute develops between the patient and what British neurologist MacDonald Critchley (1900–1997) labeled the "hemiplegic twin." It seems only a small step from phantom hemibody to full-blown reduplications of one's entire body, such as in *doppelgänger* and out-of-body experiences. The study of these neuropsychiatric aspects of corporeal awareness and their possible evolution from the relatively simple case of the phantom limb may in the future lead to fundamental insights into how the brain mediates our experience of an embodied self.

Peter Brugger

See also Body Perception; Body Perception: Disorders; Cortical Reorganization Following Damage; Experience-Dependent Plasticity; Out-of-Body Experience; Pain: Neuromatrix Theory; Proprioception

Further Readings

Brugger, P. (2006). From phantom limb to phantom body. Varieties of extracorporeal awareness. In G. Knoblich, I. Thornton, M. Grosjean, & M. Shiffrar (Eds.), *Perception of the human body from the inside out* (pp. 171–209). Oxford, UK: Oxford University Press.

Brugger, P., Kollias, S. S., Müri, R., Crelier, G., Hepp-Reymond, M.-C., & Regard, M. (2000). Beyond re-membering: Phantom sensations of congenitally absent limbs. *Proceedings of the National Academy of Sciences USA, 97,* 6167–6172.

Flor, H., Nikolajsen, L., & Staehelin Jensen, T. (2006). Phantom limb pain: A case of maladaptive CNS plasticity? *Nature Reviews Neuroscience, 7*(11), 873–881.

Giummarra, M. J., Gibson, S. J., Georgiou-Karistianis, N., & Bradshaw, J. L. (2007). Central mechanisms in phantom limb perception: The past, present and future. *Brain Research Reviews, 54,* 219–232.

Melzack, R. (1992). Phantom limbs. *Scientific American, 266,* 120–126.

Price, D. B., & Twombly, N. J. (1976). *The phantom limb phenomenon: A medical, folkloric, and historical study.* Washington, DC: Georgetown University Press.

Price, E. H. (2006). A critical review of congenital phantom limb cases and a developmental theory for the basis of body image. *Consciousness and Cognition, 15,* 310–322.

Ramachandran, V. S., & Hirstein, W. (1998). The perception of phantom limbs: The D.O. Hebb Lecture. *Brain, 121,* 1603–1630.

Wade, N. (in press). Beyond body experiences: Phantom limbs, pain and the locus of sensation. *Cortex.*

PHENOMENOLOGY (PHILOSOPHY)

As a branch of philosophy, *phenomenology* is the study of consciousness and the objects of focus within our private mental experience. In this entry, Edmund Husserl, the founder of phenomenology, will be discussed, along with his connection to early experimental psychology, as well as intentionality, bracketing, and direct reference as they relate to phenomenology. As psychology evolved into an autonomous science during the late 19th century, its agenda was steeped less in historical philosophical dogmas and agendas and more in the exploration of conscious experience and its attributes. *Phenomena* are the direct, central objects of attention. Our awareness of phenomena is instantaneous, even if their meaning is not. These raw, "unprocessed" mental events occur without analysis, opinion, or judgment. Indeed, the phenomenological approach to conscious experience requires the suspension of personal habits of thought, memory, and cultural influences. Whether this is possible has been debated vigorously. Phenomenology holds a strong nativist perspective, as opposed to a theory of mind requiring cognitive synthesis of "elements" of mental life. For the phenomenologist, immediate experience confirms timeless mental facts.

Edmund Husserl

Edmund Husserl (1859–1938) was a student of both Franz Brentano at the University of Vienna and Carl Stumpf at the University of Halle. The publication of *Logische Untersuchunger* ("Logical Investigations") in 1900 through 1901 represents the beginning of phenomenology. Husserl was distressed by the diminished importance of ancient philosophy and the growing influence of naturalism in the European scientific tradition. He believed that philosophy was moving away from what was true, immutable, and eternal and toward conceptions of truth that were tied to ever-changing, contemporary, relativistic standards. It was Husserl's deeply held belief that philosophy must address a clear understanding of what was essential and irrefutably true in nature. Phenomenology, he believed, offered a glimpse of the essence of human conscious experience, unaffected by constantly emerging (and often conflicting) scientific discoveries about the nature of man and the mind.

For Husserl, phenomenology was much more complex than an examination of subjective experience or a person's idiosyncratic perceptions of cognitive events. It was Husserl's goal to inform psychologists of the structure of mental events that supported such activities as memory, recognition, and expectancy—a more ambitious agenda than the one begun by Wilhelm Wundt (1831–1920) at Leipzig in the mid-1870s with the application of trained introspection to the elements of conscious experience. "Naturalizing consciousness" with the methodological tools and concepts of the natural sciences was, for Husserl, an error. He would never accept the premise that consciousness was a phenomenon of physical reality that followed all the laws and principles of physical science. He could not accept the identity of consciousness and psychophysiological processes in the central nervous system.

This raises the question as to the nature of these basic, background psychological events that were thought to be behind our most vivid, personal perceptions. Husserl's objective was to study the nature of "essences" in his attempt to answer these questions. In a nativistic sense, *essences* are identical for all people in all places and have been for all time. An essence is a subjectively experienced point in time with no shared, agreed-upon linguistic descriptors. It is an understanding of the supraordinate descriptors of a class of objects. For example, houses exist in many sizes, shapes, and architectural styles. They are built of brick, stone, boards, vinyl, and aluminum. Their windows have a variety of appearances. They are surrounded by beautifully landscaped lawns or malodorous concrete walkways. The phenomenologist here asserts that we comprehend (and recognize and recall) the quality of "houseness" wherever we go. Enclosed dwellings that protect inhabitants from the elements are too simplistic generalizations but approaches the idea of Husserl's "essence" of "houseness."

Husserl's phenomenology was immersed in the immediate experience of essences without further analysis or dissection of mental events, as later became apparent among structuralist psychologists.

Intentionality

For Husserl, the fundamental attribute of consciousness is *intentionality*. All functions of consciousness, either brief or prolonged, are directed toward an object. All intentions have a *direction*. Even our attention to the content of our consciousness involves intention and direction. Whenever we think, we think of something. Whenever we perceive stimuli that come from outside ourselves or inside ourselves, our perceptions have both a purpose and a direct object. Whenever we experience desire, elation, or distress, our mental contents include both focused, directed contemplation and an affective state of more or less concrete reality. For Husserl, no interpretation accompanied awareness of these contents of consciousness.

Bracketing

The method by which one attempts to briefly suspend interpretation of mental events is known as *bracketing the natural attitude.* Bracketing is a fundamental method of phenomenological inquiry. Bracketing is the attempt to focus not on the contents of thought, but instead on the basic acts of consciousness directed at those thoughts. The subject attempts to exclude from consciousness their biases, preconceptions, and all else they might think or feel about the contents of consciousness. Bracketing is also referred to as the *phenomenological reduction.* Phenomenologists are aware that bracketing can never be complete. As our self-scrutiny proceeds and preconceptions are acknowledged, there are frequently still additional ones behind them. From this perspective, human observers of their own conscious processes can never completely break with their ways of perceiving either external or internal reality.

Direct Reference

Within the phenomenological tradition, awareness of inner feelings and intuitions can occur without a person having the ability to verbally describe the nature of these private, imminent thoughts. This has been called *direct reference.* Husserl asserted that this experience is an essence, and that all human beings throughout history have had such feelings. Of course, to have something “on one’s mind” and be unable to articulate it is not unique to phenomenological approaches to mental life and has long been the focus of various psychotherapeutic methodologies. Direct reference represents an awareness of impending phenomena, without any specific cues as to their nature, attributes, or utility. The meaning of these phenomena is not yet grasped in thoughtful consciousness.

Husserl and Early Experimental Psychology

Husserl did not believe that a psychology couched in introspective methods could elucidate the fundamental characteristics of consciousness. He asserted that psychophysical research could not be fruitfully undertaken until the nature of a “perception” could be fully characterized in terms of its essences. The introspectionism of Wundt demanded that formally trained subjects report their perceptions according to predetermined categories of assessment and response and not universal, invariant acts of consciousness that could be experienced and described without training or laborious instruction. Husserl believed that phenomenology could assist psychologists in their qualitative attempts to apprehend mental contents. For him, empirical psychology could not mature without prior phenomenological description of the contents of consciousness. Such description would be devoid of a subject’s personal values or preconceptions. Phenomenology required native, holistic assessments of the objects of our attention, not analytic interpretations.

Sanford Lopater

See also Attention and Consciousness; Consciousness; Intentionality and Perception; Psychophysical Approach; Unconscious Processes

Further Readings

Hein, S. R., & Austin, W. J. (2001). Empirical and hermeneutic approaches to phenomenological research in psychology: A comparison. *Psychological Methods, 6,* 3–17.

Jennings, J. L. (1986). Husserl revisited. The forgotten distinction between psychology and phenomenology. *American Psychologist, 41,* 1231–1240.

Smith, D. W. (2007). *Husserl.* New York: Routledge.

Young, P. T. (1924). The phenomenological point of view. *Psychological Review, 31,* 288–296.

PHEROMONES

Humans seem to be fascinated by the idea that unseen chemical signals termed pheromones might be subconsciously guiding their actions without their conscious awareness. *Pheromones* are chemical signals that are sent between members of the same species. They are often effective in eliciting responses at much lower concentrations than typically required for the perception of odors and can have dramatic effects on animal behavior. It has been estimated that a male silk moth needs to detect as few as eight molecules of the sexual attractant pheromone bombykol to elicit a full behavioral response, attracting the male to the receptive female that is producing it. It is unlikely that pheromones could be having such dramatic effects within the complexities of human social interactions. However, there is the possibility that human physiology and behavior may be influenced in more subtle ways by pheromonal signals without their being perceived directly.

Although the term *pheromone* is in widespread use, among both scientists and the general public, there is no consistent definition of the term. Even in the scientific literature, different researchers apply different definitions, which often result in confusion. The most widely accepted definition was published in 1959 by Peter Karlson and Martin Lüscher, who defined a pheromone as substance secreted into the environment by an individual, which elicits a definite reaction in another individual of the same species. Initially, pheromones were used to refer to chemical signals used by insects, and there has been extensive discussion about whether the term can usefully be applied to vertebrates. In many ways, the debate comes down to the semantics of what is meant by a "definite response." Although responses to pheromones can be elicited robustly in invertebrates, such as the silk moth, responses are rarely as consistent in vertebrates, as their behavior is influenced to a greater extent by context and learning. This entry focuses on the nature of pheromones and their effects, human pheromones, pheromone production and detection, and the vomeronasal organ.

The Nature of Pheromones and Their Effects

Pheromones can be volatile organic molecules that disperse readily in air. For instance, alarm pheromones, aggregation pheromones, and sexual attraction pheromones need to be volatile to convey information over a distance. However, pheromones can also be nonvolatile substances, such as peptides or proteins, that are used when information needs to be linked to an individual producer or a specific location. For example, mouse urinary proteins convey information about the sex and individual identity of territorial urine marks, which can be used by females to assess the competitive fitness of potential mates. Pheromones are not just used by terrestrial species, and the relevant parameter for the aquatic environment is solubility rather than volatility. Pheromones can be single molecules, but more often they are most effective as mixtures of molecules, known as pheromonal blends. Traditionally, pheromones have been classified into releaser and primer types, although individual molecules can convey both releaser and primer effects.

Releaser Pheromones

Chemical signals that elicit a specific and immediate behavioral effect are known as releaser pheromones. One of the best mammalian examples is the rabbit mammary pheromone. Rabbits have an extreme form of maternal care, which involves suckling their young for a single period of 3 to 5 minutes once a day. The offsprings' survival depends on a pheromone, which elicits a stereotyped nipple search behavior that guides the pups to their mother's nipples. The rabbit mammary pheromone has been identified as 2-methylbut-2-enal, which can elicit full nipple search behavior even when presented on a glass rod, outside the normal suckling context. Another example of a mammalian releaser pheromone is the boar sexual attractant, which is a blend of androstenone and androstenol. These androgen derivatives are released into the air from the saliva of sexually aroused boars and elicit attraction and adoption of the mating stance in receptive sows.

Primer Pheromones

Pheromones that elicit longer-term effects on the endocrine state or development are termed primer pheromones. For example, testosterone-dependent volatiles in male mouse urine, such as (*S*)-2-*sec*-butyl-4, 5-dihydrothiazole, (*R*, *R*)-3,4-dehydro-*exo*-brevicomin, and alpha and beta farnesenes, accelerate puberty in prepubertal female mice and induce estrus cycles in adult female mice.

As researchers' understanding of the use of chemical signals by vertebrates has advanced, new classes have been identified that do not meet the original definition of a pheromone and have led some researchers to propose new categories of pheromonal effect. The term *signaler pheromone* has been proposed for chemosignals that convey information that biases behavioral choices without directly eliciting them. This would include individuality chemosignals, such as major histocompatibility associated peptides and major urinary proteins that can influence mate recognition in mice. Chemosignals that potentially alter mood in humans have been described as *modulator pheromones*. However, these new pheromonal classifications are not as widely accepted as the original classifications of releaser and primer pheromonal effects.

Pheromones and Learning

Nonpheromonal odors can become capable of eliciting the same responses as innate pheromones following learning. For example, exposure of rabbit pups to an artificial odor paired with the rabbit mammary pheromone, in the absence of suckling, will condition the full nipple search response to the paired odor when it is subsequently presented alone. Some pheromones appear to be innately rewarding, such as the nonvolatile protein constituents of male mouse urine, which are innately attracting to female mice. These not only elicit investigation, but will also cause conditioning of volatile urinary odors, which increases their subsequent attractiveness in the absence of the nonvolatile pheromonal signal. The rewarding effects of such pheromones result in the learning of associated contextual cues that reinforce the innate action of the pheromone.

Human Pheromones

Is human behavior too complex to be governed by simple pheromonal cues? Humans are certainly smelly primates and have axillary secretions that potentially provide a rich source of putative pheromonal signals. However, many years of scientific research have so far failed to identify robust behavioral effects mediated by such secretions. This does not necessarily mean that chemical signaling isn't able to exert subtle effects over human behavior, but rather that the complexities of human society may make it difficult to identify such effects. Genomic analysis reveals a general decline in olfactory function in humans, which has accompanied the development of visual and verbal signals and complex social interactions. Of the approximately 960 genes for classical olfactory receptors in humans, over 51% are nonfunctional pseudogenes. Nevertheless, chemicals present in human axillary secretions have been found to cause hormonal changes in recipients. For example androstadienone, a compound present in male sweat, has been shown to increase levels of the hormone cortisol in females. Axillary secretions from females have also been found to influence menstrual cyclicity of recipient females, suggesting that they contain pheromonal substances, which may underlie synchronization of menstrual cycles in female communities. Although it is difficult to demonstrate convincing pheromonal effects on adult human behavior, the search for pheromones influencing mother–infant interactions may prove more fruitful. For instance, Montgomery's glands in the areolar region around the nipple produce a milky secretion, which has been suggested to contain a human equivalent of the rabbit mammary pheromone that facilitates suckling.

Pheromone Production and Detection

Different species employ a great variety of sources of pheromones and mechanisms for their release into the environment. In many cases, pheromones are released in excreta, such as urine and feces, and are often used for territorial marking. Pheromones can also be released in secretions, such as vaginal, tear, and skin secretions, or saliva. Many species also have specialized scent glands for the release of pheromones, such as flank glands, chin glands, or sternal glands, which can be used for marking and advertising their presence. Pheromonal detection systems need to be extremely sensitive to detect the small amounts of chemical signals that are released into the environment. Pheromones may therefore act at concentrations below the threshold for conscious perception, although they may also be perceived by the main olfactory system if present in high enough concentrations. For example, heptanone acts as a male urinary pheromone that influences female reproductive states at concentrations over 100-fold lower than those required for perception. However, even when pheromones may not be perceived as odors, any autonomic and endocrine changes they elicit could affect perception of body state.

Although recent advances in molecular biological techniques and comparative genomics have advanced our understanding of the sensory systems underlying pheromonal communication in mice, our understanding of pheromonal communication in other species, including humans, is still rudimentary. Many well-known pheromonal effects, such as the rabbit mammary pheromone and the boar sex attractant pheromone, are mediated by the main olfactory epithelium. In addition to the classical olfactory receptors, subpopulations of sensory neurons of the main olfactory system have been found to express other families of chemosensory receptors. For example, a subpopulation of sensory neurons expresses trace amine associated receptors and responds to volatile amines, such as beta-phenylethylamine, which has been associated with stress responses in mice and humans. A further subpopulation of mouse olfactory sensory neurons express guanylyl cyclase receptors and form a chemosensory subsystem that appears to respond to certain urinary peptides, although whether or not these convey pheromonal effects remains to be determined.

In addition to the main olfactory epithelium, several specialized chemosensory structures are present in the nose that could mediate pheromonal effects. For example, the Grueneberg ganglion is found at the tip of the nose in mice and contains chemosensory cells that sense alarm pheromones in mice, which elicit aversion and freezing behavior. The Grueneberg ganglion has also been found in humans, although whether it has a similar role in mediating stress and fear responses in humans remains to be determined.

Vomeronasal Organ

The vomeronasal organ is a blind-ended tubular structure located in the nasal septum with a duct opening into the nasal and/or oral cavities. It is a major sensory organ for pheromonal detection, although it is also capable of detecting nonpheromonal chemosignals from potential predators and prey. Nonvolatile chemosignals are pumped into the vomeronasal organ following direct physical contact with a scent source, such as a urine mark. This is achieved by a vascular pumping mechanism in rodents, but in other species involves the pressure changes caused by curling of the upper lip, known as flehmen behavior. Broadly speaking, the vomeronasal system is specialized for the detection of nonvolatile chemosignals, in contrast to the main olfactory system that responds to volatile odorants. However, there is considerable overlap between the two systems, and studies at the receptor level have shown that the main olfactory and vomeronasal sensory neurons can each respond to both volatile and nonvolatile chemosignals. The vomeronasal system forms a more direct and specialized neural pathway for pheromonal effects, although the main olfactory system can also mediate pheromonal responses via its direct projections to hypothalamic centers controlling neuroendocrine and behavioral output.

There are two classes of vomeronasal receptors. The V1r class has over 150 members in mice, which respond to small volatile chemosignals, such as the testosterone-dependent pheromones in male mouse urine. These V1r mediated pheromonal signals elicit aggressive behavior by male mice and by maternal females toward rival adult males. The V2r class of vomeronasal receptors contains around 60 functional receptors in mice. They respond to peptide and protein chemosignals that convey signals of individual identity, which are important for territorial behavior and social recognition. However, the V1r and V2r gene families show probably the greatest rate of evolution of any vertebrate genes, leading to major species variations. Many species lack functional V2rs entirely, and many others have a far more restricted range of V1rs, suggesting that pheromonal signaling in rodents should not be assumed to be typical of other mammalian species.

A significant feature of vomeronasal receptors is their high sensitivity. V1r receptors typically respond to volatile pheromones, such as alpha and beta farnesenes, at concentrations of around 10^{-8} moles per liter. V2rs are even more sensitive and can respond to major histocompatibility associated peptides at the astonishingly low concentration of 10^{-13} moles per liter. Vomeronasal receptors also typically have more selective responses than olfactory receptors, and fail to show a broadening of their response profiles as the stimulus concentration is increased. This supports the view that the vomeronasal system may utilize labeled line sensory coding rather than the pattern recognition coding of the main olfactory system.

Although a vestige of the vomeronasal organ may be found in the nasal septum of some humans, it lacks a developed sensory epithelium and nerves to connect it to the brain. Moreover, almost all of the genes for vomeronasal receptors and transduction mechanisms are pseudogenes in humans. A functional human V1r-like gene has been identified in the human genome, but the receptor is expressed in the main olfactory epithelium, and whether it has any function in pheromonal communication is unknown. Therefore, pheromonal effects in humans are likely to be mediated via receptors in the main olfactory epithelium, such as those expressing trace amine associated receptors or possibly via the Grueneberg ganglion in the case of alarm pheromones.

Peter A. Brennan

See also Animal Chemical Sensitivity; Olfaction; Olfaction and Reproductive Behavior; Olfactory Receptors and Transduction; Vomeronasal System

Further Readings

Brennan, P. A., & Zufall, F. (2006). Pheromonal communication in vertebrates. *Nature, 444,* 308–315.

Hurst, J., & Beynon, R. (2004). Scent wars: The chemobiology of competitive signalling in mice. *Bioessays, 26,* 1288–1298.

Keller, A., & Vosshall, L. B. (2008). Better smelling through genetics: Mammalian odor perception. *Current Opinion in Neurobiology, 18,* 364–369.

Leinders-Zufall, T., Brennan, P., Widmayer, P., Chandramani, P. S., Maul-Pavicic, A., Jäger, M., et al. (2004). MHC class I peptides as chemosensory signals in the vomeronasal organ. *Science, 306,* 1033–1037.

Liberles, S. D., & Buck, L. B. (2006). A second class of chemosensory receptors in the olfactory epithelium. *Nature, 442,* 645–650.

Wyart, C., Webster, W. W., Chen, J. H., Wilson, S. R., McClary, A., Khan, R. M., et al. (2007). Smelling a single component of male sweat alters levels of cortisol in women. *Journal of Neuroscience, 27,* 1261–1265.

Wyatt, T. D. (2003). *Pheromones and animal behaviour.* Cambridge, UK: Cambridge University Press.

PHILOSOPHICAL APPROACHES

Philosophers think of perception as a sensory experience (also called a *sensation*) that provides the perceiver with knowledge concerning the external world. This leads into an approach to perception that diverges importantly from those taken in science and in everyday thinking.

Three clarifications:

1. A *sensory experience* is one that has a characteristically visual, auditory, or other-sensory character. Such an experience need not originate from the sense organs.
2. "External" means external to the conscious mind. Accordingly, the perceiver's own body should be considered a part of the "external" world.
3. In philosophy, "knowledge" means something quite specific. In order for me to *know* something in the philosophical sense—for instance, for me to *know* that my friend is walking up the path to my front door—it should not only be *true* that she is doing this, but, more than this, I should be in a position to *guarantee* that she is. (A guarantee need not be absolute; many consider it enough if one can show that something is overwhelmingly likely.)

Unpacking the previous conception, then, perception is a sensory experience that allows me to guarantee some condition of the external world. Suppose I look out of the window and seem to see my friend walking up the path. Does this give me what I need to guarantee that she is? Only if it does, does my experience count as visually perceiving her. (This will be qualified and clarified in the following section.)

Notice that many sensations do not even begin to count as perception by this criterion. Somebody hits you on the head, and you "see stars." You are in a very quiet room and "hear" internally generated white noise. You have been sitting in an airplane seat too long and feel "pins and needles" in your feet. In these cases, you have experiences with visual, auditory, or tactual character. But these experiences do not allow you to guarantee anything about the external world. They do not amount, therefore, to perception. In this entry, the philosophical approaches to perception are covered, including direct perception (including skepticism), sense data, simple and complex qualities, representationalism, and the theories of representational content.

Direct Perception

Now let's take a look at how philosophers put the concept of direct perception to work. You see a familiar-looking object—a chair perhaps—and are asked how far away it is. Without hesitation, you respond: "10 feet" or "3 paces." Given normal vision, humans can make such estimates quickly and accurately. Animals, too, are good at this kind of task: Of course, they cannot verbally answer the question, but they seem to sense how far away things are and adjust their behavior accordingly. For example, a dog might just stretch its paw for a nearby morsel of food, but if the food is far away, it would get up and walk to it. This shows that it can estimate distance. From these familiar facts, a scientist might conclude:

I. We perceive distance by vision.

And this is indeed what we would say in everyday contexts.

But wait! Suppose that you are viewing the chair through a peephole in a psychology lab. It *looks* 10 feet away, but the experimenter might be playing a trick on your eyes. The chair might be much larger than the ones you normally encounter, and because of the odd perspective forced on you by the peephole, you may have failed to appreciate its true size. Though it *looks* 10 feet away, it may actually be 15 feet away. Can you be *sure* that this is not what is going on? Maybe you should be less confident. Maybe you can't *guarantee* that it is 10 feet away.

Of course, this is a special situation. When you are walking around your own living room, you can be confident that no such visual tricks are being played. What is the difference? Reflecting on this, we see that the path from sensory experience to knowledge goes something like this:

A. The chair *looks* 10 feet away.

B. The viewing conditions are normal (in terms of perspective, lighting, etc.).

C. Therefore, the chair is really 10 feet away.

A classic philosophical position is that a sensory experience, exampled here by A, leads to knowledge about distance, as in C, only when taken together with additional *non*perceptual beliefs, such as B. The difference between living room and lab lies in the level of confidence you attach to B.

This reasoning leads us to the following conclusion: Visual experience does not give us knowledge of distance *by itself*. In view of the conception of perception introduced earlier, many philosophers put this as follows:

II. We do not *directly* perceive distance by vision.

By a consideration of the conditions under which perception leads to knowledge, thesis II introduces a notion of *direct perception* that was absent in the reasoning leading up to thesis I. We *directly* perceive something if we have a sensation that guarantees it without the need for additional information.

Skepticism and Direct Perception

Now, let us push the argument a bit further. We have been assuming that there is something in common between a *true* visual experience of a certain situation (the chair looking 10 feet away in your living room) and a *false* one (a similarly shaped, but larger chair misleadingly looking the same distance away in the lab). The experiences are the same, but one is true and the other false.

Now, one can imagine that for *every* true sensory experience, there is a similar false one. Philosophers have often used the device of a dream to bring this point home. Could you not be dreaming right now? Maybe there is no book in front of you, no entry on philosophical approaches to perception—maybe this is just a dream. Maybe your sensory experience right now is just like what you would have if there really was a book in front of you—only in this case, you are dreaming and it is false. In a similar vein, the popular movie *The Matrix* envisages a whole population whose experiences are induced by computer-controlled brain stimulation. Each of these people thinks she or he is living a life just like you imagine yours to be—but actually, she or he is just a body in a bath of chemical nutrients, hooked up to a computer that feeds the brain a rich and coherent, but wholly empty, stream of experiences. They think that they do normal things, like going to the office or to discos, seeing a beautiful woman in a red dress, or getting abused by the police, but all these experiences are simply fed to them by the computer. The question is: What makes *you* (the reader) so sure that your experience of this book is not empty—just like those in *The Matrix*?

The previous paragraph is known as a *skeptical* argument—*skepticism* is the position that one never knows anything.

Again, one may contrast two situations here. In the "normal" one, you have the perceptual experiences you are now having, and they are (mostly) true. In *The Matrix* situation, you have exactly the same experiences, but they are all false. These perceptions are similar from your point of view. So it might seem, just as in the argument A through C in the preceding section, that the experience will not by itself guarantee knowledge about the external world. We need a premise like B over and above the experience, something like, "This is a normal situation (not a *Matrix* situation)."

Some philosophers take just this position, but many others, perhaps a majority, do not want to go down this path. *Matrix*-type situations are too implausible to pose much of a threat to the guarantees you get from perception, they reason. (Remember that most philosophers do not insist on the knowledge guarantee being absolute.) They acknowledge that B is needed to rule out the eye-deceiving situations that occur in everyday life, but they do not agree that a corresponding premise is needed to rule out skeptical situations.

Consider, however, the following argument:

A. If two experiences are exactly similar (from the inside, as it were), then whatever one directly perceives in one, one also directly perceives in the other.

The idea behind D is that if a perceiver cannot distinguish between two experiences, then they must be equivalent. (D is controversial, as we shall see, because there are other ways of fleshing out this equivalency.) Continuing with the argument:

B. One's experiences in the normal and *Matrix* situations *are* exactly similar.

C. In the *Matrix* situation, there is nothing external that one directly perceives.

D. Therefore, in the normal situation, there is nothing external that one directly perceives.

Like A through C, this argument compares a true perception with a false one. But, courtesy of proposition D, it has a very radical conclusion:

III. We never directly perceive anything external.

This poses a big problem for the theorist because it makes it difficult to figure out how we gain knowledge of the external world by perception.

Sense Data

Here, some philosophers avail themselves of a clever move that traces back to the great 17th century philosopher and mathematician, René Descartes (hailed by some as the founder of modern philosophy). Descartes propounded the skeptical situation we have just discussed (in its low-tech version, of course—the dream). But he noticed that these situations still leave room for certainty about *something*. You are not entitled to be certain that there is a book in front of you, Descartes argued, but you can be certain about something else, namely that *it looks and feels to you as if* there is a book in front of you. This is not certainty about an *external* fact, but there is certainty about what your experience tells you about the external world, putting aside the question of whether what your experience tells you is true or false.

Back to direct perception. Suppose that you hold that in every perceptual experience, one directly perceives *something*. As we have seen, thesis III leaves you with a problem, for it implies that you do not directly perceive external things such as chairs, tables, and books. Descartes' move helps here. For we can now hold that every sensory experience is a direct perception of things like the *look* of tables, the *feel* of books—more generally, of a *perceptual appearance*. (As seen later in the Representationalism section, this is a controversial move—nonetheless, it was very influential in its time.) The look of a table can in principle give you knowledge about the table itself (i.e., about the external world, though it is controversial exactly how).

Philosophers call perceptual appearances *sense data*. Up until roughly 1960, most philosophers held that:

IV. All perceptual experiences are direct perceptions of sense data.

The claim is that what we directly perceive in both the lab and living room (see previous paragraphs) is a-chair-ten-feet-away sense datum, and in both the normal and the *Matrix* case, a book-in-front-of-me sense datum. The cases in each pair differ with respect to what is present in the real world, but in accordance with D they are similar with respect to what is directly perceived.

Simple and Complex Qualities

There is one other important consideration regarding what we directly perceive. The senses engage directly and automatically with certain features of the environment. For example, vision tells us about color and shape without any assistance from any other faculty. We don't *learn* to perceive color (or the appearance of color); color vision instinctually emerges in infants without the need for any teaching or prior experience. Color is not known through anything else, and for this reason, one will not fully appreciate what it is unless one experiences it. The same is true for pitch, timbre, and loudness in audition, and of various tactile qualities. We call qualities like color *simple.*

Things are different when, for instance, we visually recognize something as a daffodil or an oak. Nothing appears to the unknowledgeable viewer as a daffodil; he must learn, or be taught, that certain visual features (color, shape, etc.) go together in a particular type of flower. One grasps the visual idea of a daffodil through knowing how these other visual features go together; thus, it is possible (in principle at least) to understand what a daffodil looks like without ever having seen one—somebody could just tell you. (But they cannot get you to grasp what *yellow* looks like simply by telling you.) Ideas of this sort are called *complex*.

Complex ideas are known through simple ones. For this reason, direct perception is usually restricted to simple ideas (or appearances thereof). Strictly speaking, therefore, there is no *chair* sense datum or *book* sense datum, as asserted at the end of the preceding section. These are properly specified in terms of *simple* sense data, such as color and shape.

Representationalism

Thesis IV was widely held for a long time. But it beggars common sense by supposing that there are *appearances* over and above material objects and their real qualities. (For sense data are acknowledged to be immaterial entities.) Most philosophers today feel that such a notion is unsustainable. There is something unnatural in the supposition that we directly perceive immaterial things. It is far more natural to hold that perception relates us to material objects and the qualities that these objects have independent of perceivers. This is of a piece with thesis I, that humans and other animals perceive distance. The challenge is to construct a philosophically viable framework for such claims in light of the conception of perception as contributing to knowledge.

How did sense datum theory get into the unattractive position of positing immaterial appearances? By substituting appearances for the things we perceive. Common sense says that we perceive material objects. The sense datum theorist contests this. She argues that perception is *not* a relation between perceivers and material objects. Rather, it is a relation between the perceiver and immaterial sense data.

The philosophical position known as representationalism attempts to deflect this attack on common sense. It is based on a subtle reinterpretation of Descartes' insight. What does it mean to say that it *appears* as if something is so—for instance, that it *looks* as if there is a blue sphere 10 feet away, or that it *feels* as if the surface at one's fingertips is rough, or that it *sounds* as if there is a constant low buzz in the room? According to the representationalists, such statements specify the *message* conveyed by a particular perceptual experience—a message that is *true* if there really is a blue sphere 10 feet away, or if the surface is indeed rough, and *false* otherwise.

One could put it in this way: The representationalist conceives of perceptual experiences as akin to messages spoken by an informant. The informant *tells* you, for instance, that Home University (HU) has won the game. Your knowing that she has told you this is different from knowing that HU has really won—after all, she might be mistaken or even be lying. Knowing what she says is not a matter of knowing whether HU has won the game or not; rather, it is a matter of knowing *what the world would be like if what she said was true.*

Similarly, an auditory experience "tells" you that a certain kind of sound is present—for instance, that there is a constant low buzz in the room. When you have a perceptual experience, you instinctively understand the "message" it conveys. You know what the world *would* be like, *if* your perception were true. For instance, you know, without having to learn it, that a particular kind of auditory experience tells you of a constant low buzz, while another auditory experience, of a different character, tells you of a sudden, loud, high-pitched scream. Descartes' insight can be recast in this way: Knowing that your senses tell you of some occurrence is much less problematic than knowing that what they tell you is true.

Here then is the difference between the sense datum theorist and the representationalist. Both agree, following Descartes, that we know how things look but may be mistaken about how things really are. But they give different accounts of what this means. According to the sense datum theorist:

> It looks to me as if there is a blue sphere there = I see a blue-sphere-there sense datum.

According to the representationalist, by contrast:

> It looks to me as if there is a blue sphere there = If my current visual state is true, there *is* a blue sphere there.

Notice that the representationalist is *not* committed to saying that if two experiences are indistinguishable from the perceiver's point of view, then they are direct perceptions of the same thing. (Compare D above, in the sense datum theorist's argument.) A false perception may be indistinguishable from a true one; yet, there is no blue sphere to be seen in the case of a false perception.

Theories of Representational Content

The "message" conveyed by a perceptual state is known as its "content." Content is known instinctively—one knows immediately upon experiencing it that a visual state announces the presence of something blue or something spherical. But what property in the external world does the blue sensory state represent? This question is difficult to answer because the sensory state does not seem to give us perceiver-independent information about color. We have no instinctive way of going from the *experience* of blue to the physical character of blue things. Taken by itself, our experience gives us no hint, for example, that color is closely related to wavelength. Yet, color experience does presumably represent a wavelength-related property.

Causal theories of representation try to say what experience represents. Roughly:

> A perceptual experience represents property P in virtue of the fact that P causes that experience.

Thus, experiences of blue represent a certain wavelength-related property because occurrences of this property cause these experiences to occur. Causal theories run into an immediate difficulty. They find it difficult to accommodate content in the case of false or mistaken perceptual states. A perceptual experience of blue might, in certain circumstances, be caused by something green. Why does it represent *blue* in this case?

One attempt to address this difficulty in causal theories of content is the so-called teleosemantic approach. This proposes that the content of a perceptual experience is that to which it is tied by biological function. (Functions are result-oriented, or "teleological," hence the name.) The *function* of a blue-announcing experience is for it to occur when a blue thing is in view, and thus to ready the perceiver for this fact. But because things can sometimes malfunction, nonblue things can sometimes cause such an experience. Nonetheless, if one can assume that visual systems function reasonably well, accuracy is a good bet. This explains why it is reasonable to rely on perception.

Other theories abandon the environmental focus of cause and redirect their gaze downstream to what the organism instinctively *does* with perception. *Sensorimotor theories* are one attempt to do this; they claim that perceptual experience represents the behavior that *follows* from perception, not the world outside. Another radical approach is "direct realism," according to which, perception is a direct engagement with the world with no representational intermediary, whether in terms of sense data or in terms of representational content. Yet another approach is to accept the notion of content, but to treat of it "externally" (i.e., as deriving its significance from the external things with which we interact, rather than to a purely subjective image of these things). All of these approaches tend to deny the input-output separation of an organism's engagement with the world, preferring to view action and perception as two aspects of the same complex process.

Conclusion

Philosophers tend to take an attitude toward perception that is quite unfamiliar to psychologists and neuroscientists, and for that matter, to artists and ordinary people. A recent, well-received philosophical work on perception, *The Problem of Perception* by A. D. Smith, proclaims on page 1 that "when philosophers speak of 'the' problem of perception, what they generally have in mind is *the question whether we can ever directly perceive the physical world*" (emphasis added). At first sight, such a question seems bizarre. Hopefully, the reader has now begun to glimpse why such a question is asked, and how it is addressed.

Mohan Matthen

See also Action and Vision; Causality; Color: Philosophical Issues; Content of Perceptual Experience; Indirect Nature of Perception; Intentionality and Perception; Naïve Realism; Qualia; Seeing as

Further Readings

Ayer, A. J. (1962). *The problem of knowledge.* Baltimore: Penguin.

Dretske, F. I. (1981). *Knowledge and the flow of information.* Cambridge: MIT Press.

Mackie, J. L. (1976). *Problems from Locke.* Oxford, UK: Clarendon Press.

Matthen, M. (2005). *Seeing, doing, and knowing: A philosophical theory of sense perception.* Oxford, UK: Clarendon Press.

Noe, A. (2004). *Action in perception.* Cambridge: MIT Press.

Siegel, S. (2005). The contents of perception. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy.* Stanford, CA: Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University. Retrieved from http://plato.stanford.edu/entries/perception-contents

Smith, A. D. (2002). *The problem of perception.* Cambridge, MA: Harvard University Press.

PHILOSOPHY: ACCESS AND REPORT

Perception gives us access to the world around us insofar as it enables us to register and use information that we receive from our environment. Linguistic reports are just one way in which information can be registered and used. The registration and use of information that constitutes perception requires something *more* than the eliciting of a distinctive bodily change in response to a particular feature in the environment; a nearby magnetic field may have a distinctive effect on a person's body without that person perceiving the magnetic field, and a slight shift in the vibrations of a plucked string may result in a corresponding shift in the vibrations of an eardrum without that shift being perceived. The registration and use of information requires something *less* than reporting in a language or using the information in reasoning; however, a person can perceive the contours of a tree without ever describing those contours in words (to others or to oneself), and one may notice the color of a stormy sky without ever engaging in reasoning about the sky. Perceptual access, then, requires something more than bodily sensitivity to information in the environment, but something less than linguistic reports and reasoning. This entry covers access and accessibility, conscious access, perceptual access and integration, and direct versus indirect responses.

Access and Accessibility

We can distinguish between information that is registered in such a way that it *could* affect our reasoning or our reports, and information that is registered in such a way that it *could not* affect anything we think or say. There are some sounds and some shifts in lighting that we do not report on or reason about *in fact,* but that we *could* report on or reason about if we really needed to (the distant buzz of a mosquito, for example); and it may be that it is our *capacity* for such reporting and reasoning rather than their actuality that is crucial for perception.

It is not clear, however, when a given prompt activates an already existing capacity and when it creates something new. When does the question "Do you hear a mosquito?" reveal what is already perceived, and when does it focus one's attention so as to create a new perception? If people need extensive training in order to report and act on their physiological responses to very high sounds, has this training activated an already existent capacity or has it created a new capacity and thus a new perception? It may be possible to provide some guidance about what sorts of prompts reveal rather than create the capacities that constitute perception. Open-ended questioning is more likely to reveal already existent perceptions, whereas more pointed questioning is more likely to create new perceptions, for example. Responses that flow easily are more likely to tell us what was already perceived than responses over which we struggle. But there is a lingering suspicion that already existing perceptions depend on already existing access of some sort; that it is this pre-reflective, pre-linguistic sort of access that is crucial for perception.

Conscious Access

Many people have thought that the sort of registration needed for perception is *conscious* registration: Information that "enters" our consciousness is perceived, whereas information that we are not aware of is not perceived. This suggestion would rule out the possibility of unconscious perceptions—cases in which it makes sense to say of a person (on the basis of her behavior) that she hears a buzzing sound even though she denies hearing anything. Ordinary talk is often conflicted about the need for consciousness. We say "she must have been conscious of the buzzing, at some level" but also "if we aren't aware of it, we can't hear it." The deeper problem with understanding perceptual access as equivalent to conscious access is that the initial puzzle about perception—a puzzle about what more than bodily sensitivity, but less than verbal report, is required—is now replaced by precisely the same puzzle about consciousness. Whether one insists that all perception is conscious perception, or not, one is still left with the challenge of determining the nature of something that falls "between" the mere reception of information and the active utilization of information in speech and in action.

Perceptual Access and Integration

Before any information can be registered in a way that can guide behavior, that information must be integrated with various other sorts of information concerning our environment. The received information must be situated within the spatiotemporal world in which we live, and that depends on bringing separate bits of information into relation with each other.

Registering a series of sounds as belonging to a single voice, or recognizing that different appearances are appearances of the same object, depends on a capacity to integrate disparate information into a coherent whole. Information about color and shape and textures must be integrated in such a way as to register them as properties of a single object—an object that can be acted on. This process is sometimes called *binding,* and it is mirrored in the subject–predicate groupings of the sentences we use to describe the world around us.

Information about a particular object and its properties must, in turn, be integrated with information about other objects (including oneself) and their properties before it can be used to guide action. Before we can do anything about a buzzing sound, for example, we need to locate it in space, and locating it in space depends on locating it in relation to ourselves and various other objects in the environment. This is sometimes referred to as the creation of a *mental map*.

The integrative processes of binding and the creation of mental maps seem to be essential to perception. We don't perceive anything at all until we perceive something as located at some place; and we don't perceive something as located at some place unless we recognize that it can present different aspects from different angles, standing in various spatiotemporal relations to various other objects. The perception of a cup, for example, is always the perception of a cup at a particular location; and it is not possible to perceive the cup as "there" without registering its relation to "here" and to various other "theres." Recognizing the location of objects and their properties in relation to one another is, then, precisely what seems to be required for the transition from mere sensitivity to information to the perception of what is around us.

Direct Versus Indirect Responses

Some of our responses to what we perceive are direct. Hearing an approaching mosquito, one might step to the side, swat it, or say "that was a mosquito" or "what an annoyance." Other responses are more indirect. A person suddenly becomes nervous as the buzzing approaches or suddenly begins to speak about blood. There are innumerable ways, well known to therapists and literary critics, that our perceptions are indirectly revealed by what we do and say.

Whether or not one insists that perception always includes a *capacity* for more explicit action and more explicit reporting, and whether or not one allows that some perception is *unconscious,* it seems clear that perceptions always involve the integration of information into a unified view of a spatiotemporal world. Before treating someone's sudden nervousness as indirect evidence that the mosquito was indeed perceived, there must be reason to suppose that the nervousness is not a mere reflex, that it is directed toward the mosquito as an object in one's environment. Thus, perception depends on the integration of incoming information into a larger picture of one's spatiotemporal environment.

Jennifer Church

See also Phenomenology (Philosophy); Philosophical Approaches

Further Readings

Bermudez, J. L. (2003). *Thinking without words.* Oxford, UK: Oxford University Press.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences, 18*(2), 227–247.

Church, J. (2000). "Seeing as" and the double bind of consciousness. *Journal of Consciousness Studies,* 7(8/9), 99–111.

Dennett, D. (1991). *Consciousness explained.* Boston: Little, Brown.

Dretske, F. (1981). *Knowledge and the flow of information.* Cambridge: MIT Press.

Noe, A. (2003). On what we see. *Pacific Philosophical Quarterly, 83,* 57–80.

PHILOSOPHY: ATTENTION AND THE SIZE OF THE CONSCIOUS FIELD

Several philosophical views of conscious experience have emerged. Two such views are the *rich* view and the *thin* view. On a rich view of conscious experience, there is constant experience in several modalities; for example, you have constant tactile experience of your shirt on your back, constant auditory experience of the background rumble of traffic, constant visual experience of the tip of your nose. On a *thin* view of consciousness, there is no experience without attention; that is, when you aren't paying attention to such things, they drop out of consciousness entirely, so that they form no part of your stream of experience—not even vaguely, peripherally, amorphously—no part of your phenomenology, no part of what it's like to be you.

There is, of course, perceptual *processing* without attention. A gentle tug on the shirt or an unexpected movement in the visual periphery will generally get your attention, even if you are fully absorbed in other things. To get attention, such events must first register preattentively. As your attention centers on one or a few things, you monitor many others inattentively, ready to redirect attention when an inattentional process detects a large or important change. The question is whether *experience* accompanies such inattentional perceptual processing, or whether, instead, that processing is entirely nonconscious. We might think of consciousness like a soup. Is it a rich soup, replete with experience across broad regions of several modalities simultaneously—visual experience across a broad field (though perhaps indistinct outside the central region), proprioceptive experience of the condition and position of your body, olfactory experience of the room, and perhaps emotional phenomenology, cognitive phenomenology, inner speech and imagery as well, all simultaneously? Or is experience a thin soup, limited to just one or a few regions, objects, or modalities at a time?

A clock starts chiming, but you only start attending to the chimes part way through. It seems you can count back chimes in your memory—it has chimed three times, say. What, exactly, are you remembering? Are you remembering an auditory *experience* of those three unattended chimes (as an advocate of the rich view might say)? Or are you only remembering an outward event, the chiming of the clock, that was not until now in any way experienced by you?

A variety of philosophers and psychologists have endorsed rich or thin views of experience on different grounds. (Moderate views, according to which experience outruns attention but only to a moderate degree, are also possible but less commonly endorsed.) On purely introspective grounds, William James and John Searle have endorsed rich views, whereas Julian Jaynes and David Armstrong promote thin views. Although the question might appear to be easily settled by introspection (simply introspect now and determine how much is going on in your consciousness), the divergence of opinion among philosophers and psychologists should give us pause. The central problem with concurrent introspection as a method of addressing the issue is what is sometimes called the refrigerator light illusion: The fact that you have auditory experience of the hum of traffic *when you're thinking about* whether you hear the hum of traffic provides no evidence on the question of whether you have auditory experience of the hum of traffic when you're not considering the matter. Just as the act of checking if the refrigerator light turns it on, so also might the act of checking for tactile experience of one's shirt or visual experience of one's nose produce those very experiences.

Some researchers have argued that the thin view is supported by the fact that we often fail to parse, respond to, or remember what might seem to be salient stimuli. For example, in dichotic listening tasks, people appear to recall very little about a stream of speech coming in one ear when they are instructed to attend to a different stream of speech coming in the other ear. In one striking video demonstration by Daniel Simons and Christopher Chabris, participants instructed to count the number of times one ball is passed in a complicated ball game often fail to notice a person in a gorilla suit walking through that game. Although such evidence does undermine certain versions of the rich view (according to which *detailed* representations are generated and retained without attention), it does not address rich views that allow that unattended stimuli, though experienced, are experienced in a relatively vague, inchoate, or immemorable way, or views according to which the general gist may be apprehended and remembered, but not the details.

Ned Block has emphasized that it seems introspective that we visually experience more of a visual display that we focally attend to. On the face of it, this observation seems to suggest that perceptual experience outruns attention. However, some models of attention allow for something like *diffuse* attention that spreads more broadly than focal attention, perhaps along a gradient. Consciousness might then still align neatly with attention, as advocates of the thin view believe. Even if you visually experience this whole page while focally attending only to a few words at a time, the question remains open whether you also visually experience the wall in the far periphery when you're not thinking about it, or the pressure of the shoes on your feet.

The issue of whether perceptual experience is, in general, rich or thin may also be addressed by gathering introspective or immediately retrospective reports about randomly sampled moments of experience. Eric Schwitzgebel gave people beepers to wear during ordinary activities and asked them to reflect on the presence or absence of various types of sensory experience in the last undisturbed moment before each beep. He found a majority of participants reported a visual experience in 100% of sampled moments, whereas most participants reported tactile and peripheral visual experiences in 50 to 90% of sampled moments. However, given the difficulty of distinguishing between memory of *experiences* of events from memory of outward events that were not actually experienced as they occurred (expressed previously in the case of the chiming clock), it's unclear how much confidence to put in such reports (as Schwitzgebel admits).

The size of the conscious field remains an open question, ripe for creative new arguments and research. The issue also nicely illustrates the methodological difficulties involved in the study of consciousness, due to the problems both with subjective report (whether concurrent or retrospective) and with establishing a clear relationship between conscious experience and purely objective behavioral measure.

Eric Schwitzgebel

See also Attention: Covert; Attention: Selective; Attention and Consciousness; Consciousness; Phenomenology (Philosophy); Philosophy: Access and Report

Further Readings

Armstrong, D. M. (1981). *The nature of mind, and other essays.* Ithaca, NY: Cornell University Press.

Block, N. (in press). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences.*

Dennett, D. C. (1991). *Consciousness explained.* Boston: Little, Brown.

James, W. (1980/1981). *The principles of psychology.* Cambridge, MA: Harvard University Press.

Jaynes, J. (1976). *The origin of consciousness in the breakdown of the bicameral mind.* Boston: Houghton Mifflin.

Mack, A., & Rock, I. (1998). *Inattentional blindness.* Cambridge: MIT Press.

Rensink, R. A. (2000). When good observers go bad: Change blindness, inattentional blindness, and visual experience. *Psyche, 6*(9).

Schwitzgebel, E. (2007). Do you have constant tactile experience of your feet in your shoes? Or is experience limited to what's in attention? *Journal of Consciousness Studies, 14*(3), 5–35.

Searle, J. R. (1992). *The rediscovery of the mind.* Cambridge: MIT Press.

Simons, D. J. (2000). Attentional capture and inattentional blindness. *Trends in Cognitive Sciences, 4,* 147–155.

PHYSIOLOGICAL APPROACH

Perception begins with activation of receptors in the eye, ear, skin, nose, mouth, and additional receptors inside the body. This activation results in electrical signals that are transmitted to the brain, where perception occurs. The task of the *physiological approach* to perception is to understand this process. The origins of the physiological approach can be traced back to the early philosophers and physicians. This entry covers this historical background, followed by consideration of stimuli for perception, receptor activity, and molecular processes, neural processing and neural responding, organization of sensory systems, the sensory code, making connections between physiology and perception, and how the physiological approach has been extended beyond "perception" to include recognition and action.

Historical Background

Early thinking about the physiology of the mind focused on determining the anatomical structures involved in the operation of the mind. In the 4th century B.C.E., the philosopher Aristotle (384–322 B.C.E.) stated that the heart was the seat of the mind and the soul. The Greek physician Galen (ca. 130–200 C.E.) saw human health, thoughts, and emotions as being determined by four different "spirits" flowing from the ventricles—cavities in the center of the brain. This idea was accepted all the way through the Middle Ages and into the Renaissance in the 1500s and early 1600s. In the early 1630s, the philosopher Rene Descartes, although still accepting the idea of flowing spirits, specified the pineal gland, which was thought to be located over the ventricles, as the seat of the soul.

In 1664, Thomas Willis, a physician at the University of Oxford, published a book titled *The Anatomy of the Brain,* which was based on dissections of the brains of humans, dogs, sheep, and other animals. Willis concluded that the brain was responsible for mental functioning, that different functions were located in different regions of the brain, and that disorders of the brain were disorders of chemistry. Although these conclusions were correct, details of the mechanisms involved had to await the development of new technologies that would enable researchers to more accurately observe the brain's structure and to record electrical signals in the nervous system.

By the late 1800s, researchers had shown that a wave of electricity is transmitted in groups of neurons, such as the optic nerve. To explain how these electrical signals result in different perceptions, Johannes Müeller in 1842 proposed the doctrine of specific nerve energies, which stated that our perceptions depend on "nerve energies" reaching the brain and that the specific quality we experience depends on which nerves are stimulated. Thus, he proposed that activity in the optic nerve results in seeing, activity in the auditory nerve results in hearing, and so on. By the end of the 1800s, this idea had expanded to conclude that nerves from each of these senses reach different areas of the brain. This idea of separating different functions is still a central principle of nervous system functioning.

Details about how single neurons operate had to await the development of electronic amplifiers that were powerful enough to make the extremely small electrical signals generated by individual neurons visible. In the 1920s, Edgar Adrian was able to record electrical signals from single sensory neurons, an achievement for which he was awarded the Nobel Prize in 1932.

The ability to record electrical signals from individual neurons ushered in the modern era of brain research, and in the 1950s and 1960s, development of more sophisticated electronics and the availability of computers and the electron microscope made more detailed analysis of how neurons function possible. Most of the physiological research on the senses and perception had its beginning at this point, when it became possible to determine how individual neurons respond to stimuli in the environment and how neurons work together in circuits of interconnected neurons called neural networks. In addition, the development of brain imaging technology has made it possible to determine how areas of the human brain are activated in response to specific stimuli, and a great deal of information about brain function has also been obtained from neuropsychological research, which focuses on the study of the behavior of people with brain damage.

The physiological study of the senses and perception can be described by considering a number of different processes, beginning with the stimulus and culminating in behavioral outcomes such as perception, recognition, and action. The starting point for describing how modern researchers have studied the physiology of perception is the stimulus to be perceived.

Stimuli for Perception

The perceptual process begins with a stimulus—for example, an observer looking at a tree in a field. The observer sees the tree because light is reflected from the tree into the observer's eye. Similarly, the observer hears the rustle of the tree's leaves caused by a gust of wind, because pressure changes caused by movement of the leaves travel through the air, enter the ear, and stimulate receptors inside the ear.

These examples illustrate an important principle: Stimuli in the environment, such as the tree and the sound caused by the rustle of the leaves, are perceived indirectly based on effects they have on the environment. The tree does not get into the eye and the rustling leaves do not get into the ear. The tree is represented by light reflected from it, and the rustle of the leaves by pressure changes in the air. The idea that stimuli are represented is central to the physiological approach.

Receptor Activity and Molecular Processes

The idea that stimuli are represented becomes particularly salient when we consider that everything we perceive is based on electrical signals in the nervous system. These electrical signals are created in the receptors, which transform energy from the environment into electrical signals—a process called *transduction*. Thus, the tree is initially represented by reflected light, but once transduction has occurred it is represented by electricity.

Transduction is studied mainly at the molecular level. For example, the transduction of pressure changes in the air into electricity occurs when these pressure changes are transmitted to the ear and cause structures within the ear (called cilia) to bend. This bending opens channels in the membrane of the auditory receptor that results in the flow of charged particles called ions across the membrane. This ion flow results in a series of chemical reactions that culminate in electrical signals in the receptor. A similar process of stimuli causing ion flow occurs for vision when light activates light-sensitive molecules in visual receptors and for touch when pressure activates pressure-sensitive receptors in the skin.

Receptors are important not only because they generate electrical signals, but because they help determine which aspects of stimuli are represented in the nervous system. For example, for vision, light-sensitive pigments in receptors are activated by specific wavelengths of light. These wavelengths are therefore the ones we perceive, and other wavelengths are invisible to us. For the sense of touch, there are receptors that respond best to sustained pressure, and others that respond to rapid vibration. These receptors are therefore thought to be responsible for the perception of sustained pressure and rapid vibration, respectively, and the properties of the receptors set limits on the types of tactile stimuli we perceive.

Once electrical signals are generated in the receptors, this activation is transmitted to other neurons. There are a number of different types of electrical signals. Most relevant for perception is action potential—a wave of electrical charge that travels down the length of neurons in nerve fibers or axons. Once the action potential reaches the end of the nerve fiber, it causes the release of neurotransmitter molecules at the synapse, the space between the end of the axon and another neuron.

Neural Processing and Neural Responding

Synaptic transmission is important because it enables signals to travel for long distances by jumping from one neuron to another and also because of its role in a process called neural processing. The basis of neural processing is excitation and inhibition. Excitation occurs when an excitatory transmitter is released, which increases the probability of generating an electrical signal in the next neuron. Inhibition occurs when an inhibitory transmitter is released, which decreases the probability of generating a signal in the next neuron. Each neuron receives many inputs from other neurons—over 1,000 per neuron in the brain—some excitatory, and some inhibitory. The interaction of excitation and inhibition, combined with the way neurons are connected to one another, create neural circuits that endow neurons with specialized properties.

The properties of single neurons have been characterized by determining stimuli that either increase firing rate (excitatory response) or decrease firing rate (inhibitory response). Classic experiments by Stephen Kuffler, David Hubel, and Thorsten Wiesel in the visual system and Vernon Mountcastle in the somatosensory system established the receptive field as a way of characterizing a neuron's response properties. A neuron's receptive field is the area on the receptor surface (the retina for vision, the skin for tactile sensation), which, when stimulated, influences the firing rate of the neuron.

Early research on receptive fields, which involved recording from subcortical areas and primary cortical receiving areas, discovered neurons that responded best to stimuli such as small spots of light or pressure, or to movement of bar-shaped visual or tactile stimuli in specific directions. These neurons were called feature detectors because they responded to specific stimulus features. Subsequent research, which involved recording from neurons in higher cortical areas, has revealed neurons that respond to more specialized stimuli, such as faces, buildings, and bodies (in the visual system) and to gripping specific three-dimensional shapes (in the somatosensory system).

Tuning curves are another way of characterizing neural responding. Tuning curves indicate the range of stimuli that activate a neuron. For example, an orientation tuning curve of a neuron in the visual system indicates the relationship between firing rate and stimulus orientation; a frequency tuning curve of a neuron in the auditory system indicates firing rate versus sound frequency. Generally, tuning curves indicate that neurons respond to a specific, often narrow, range of stimulus properties, although there are some neurons that respond to wide ranges of stimulus properties.

Relationships that have been established between characteristics of stimuli and physiological responding form the backbone of the idea that stimuli in the environment are represented by firing in the nervous system. Carrying this idea of representation a step further, many researchers have been concerned not only with how individual neurons respond, but with how these neurons are organized in the brain.

Organization of Sensory Systems

As previously mentioned, one of the early discoveries about sensory functioning was that neurons from each sense reach different areas of the brain. Thus, an early 20th-century conception of brain organization would locate the primary receiving areas for vision in the occipital lobe, audition in the temporal lobe, skin senses in the parietal lobe, and taste and smell in areas in the frontal lobe. Modern research has taken these primary receiving areas as a starting point, but has shown that the processing areas for the senses extend far beyond these primary receiving areas. A notable example is provided by visual stimuli, which cause activity in pathways that extend from the primary visual receiving area in the occipital lobe to the temporal and parietal lobes. Activity from other senses, such as audition and the tactile senses, also extends beyond their primary receiving areas, and activity generated by all of the senses extends into the frontal lobe, with some neurons responding multimodally to stimulation of more than one sensory modality.

Organization has also been demonstrated *within* areas of the cortex, by determining maps of various sensory functions. There are retinotopic maps for vision (a mapping of locations on the retina on the cortex, so neurons near each other in the visual cortex receive signals from nearby areas on the retina); somatotopic maps for neurons serving touch in the parietal lobe (maps of the body surface); and tonotopic maps for hearing (maps of frequency). Recent research has also revealed orderly mapping based on structural features of odorant molecules, in the olfactory bulb.

The cortex is also organized in terms of functions. Areas have been identified that are specialized to process visual information for perceiving faces, places, bodies, and motion, and auditory information for recognizing sounds and for determining their location. However, this so-called "modular" organization (although there is some controversy as to the definition of "module") should not be taken too far because although a specific type of stimulus (such as, for example, faces) might be said to be served by a specific area in the cortex called the fusiform face area, faces also cause activity in many other areas of the brain as well. Generally, therefore, specialized areas do exist, but activity for a specific stimulus is often widely distributed throughout the brain.

The Sensory Code

The idea that stimuli are represented by the firing of neurons in the brain has led to the idea of the sensory code—the specific characteristics of neural firing that are associated with specific stimuli. A number of different types of codes have been proposed: (1) specificity coding, in which a stimulus is represented by the firing of a few very specialized neurons; (2) sparse coding, in which a stimulus is represented by the pattern of firing of a small number of neurons; and (3) distributed coding, in which a stimulus is represented by the pattern of firing of a large number of neurons. Although there are neurons that are specialized to respond to a few specific stimuli, a strict specificity code is considered unlikely, both because most neurons respond to more than one specific stimulus and because the large number of possible stimuli would require a huge number of different types of neurons. There is evidence, however, for the operation of sparse and distributed coding throughout the senses.

Linking Physiology and Perception

All of the previously described research has considered how neurons or brain areas are activated by stimuli. An assumption behind this approach is that neurons or brain areas that fire in response to a specific stimulus or type of stimulus may have something to do with the perception of that stimulus. The key word in the previous sentence is "may," because until physiological processes are directly linked to perception, it is not possible to assert with certainty that the response of a particular neuron or group of neurons to, say, a face, an auditory tone, or an odor has anything to do with perceiving faces, tones, or odors.

To relate physiological function to perception, it is necessary to combine the findings from physiological research with the measurement of perception. One way this has been accomplished is using a procedure called selective adaptation. The idea behind selective adaptation is that when a stimulus with a specific property is presented, neurons tuned to that property fire, and if the neurons fire for long enough, they adapt or become fatigued. This adaptation causes two physiological effects: (1) the neuron's firing rate decreases, and (2) the neuron fires less when that stimulus is immediately presented again. According to this idea, presenting a vertical line causes neurons in the visual system that respond to vertical lines to respond, but as these presentations continue, these neurons become adapted and eventually begin to fire less to vertical lines. Adaptation is selective because only the neurons that respond to verticals or near-verticals adapt, and other neurons do not.

The basic assumption behind a *psychophysical* selective adaptation experiment is that if these adapted neurons have anything to do with perception, then adaptation of neurons that respond to verticals should result in the *perceptual* effect of becoming selectively less sensitive to verticals, but not to other orientations. Research using this technique has provided evidence linking visual orientation detectors in the cortex to the perception of orientation and by linking receptors in the skin called Pacinian corpuscles (PC) to the perception of fine textures with the fingertips. This link between the PC and fine texture perception was demonstrated by showing that adapting the PC by presenting high-frequency vibration to the fingers reduced the ability to discriminate between two fine textures with the fingertips.

A more powerful way of demonstrating connections between physiology and perception is to measure physiological responding and perception simultaneously in the same subject. This has been achieved by William Newsome in an elegant series of experiments in which he demonstrated close connections between the firing of directionally selective neurons in monkey cortex and the monkeys' perception of the direction of movement.

Another technique has taken advantage of situations in which perception changes even when stimuli on the receptors remain constant. One example of such a situation is binocular rivalry, which occurs when one stimulus is presented to the left eye and a different stimulus is presented to the right eye. In this situation, perception alternates between the two stimuli. For example, when a picture of a butterfly is presented to one eye and a picture of a starburst pattern to the other eye, the observer sees either a butterfly or a starburst, with perception alternating between the two. In experiments carried out on monkeys and humans, a specific neuron (measured using single unit recording in monkeys) or groups of neurons (measured using brain scanning in humans) respond when the observer reports seeing, say, the butterfly, but stops responding as soon as the observer reports that his or her perception has switched to the starburst. Results such as this help identify what has been called the neural correlate of consciousness, because neural firing is *correlated* with changes in consciousness, even as the stimuli on the receptors remain the same.

Beyond "Perception"

One of the central themes of the previous discussion has been that one of the goals of the physiological approach is to determine the physiological mechanisms behind perceptual experiences, such as seeing a tree in a field or hearing music at a concert. But modern research has extended beyond simply explaining *perception,* to explaining other functions such as *recognition* and *action.*

Recognition is the ability to place objects into categories—being able to identify a particular object as a "tree," a face as "Joel's face," or a sound as "John Lennon's *Imagine.*" Neuropsychological research has made clear that perceiving and recognizing are two separable abilities by demonstrating that there are people with brain damage who can see objects or hear sounds, but can't name the category to which an object or a sound belongs.

Action refers to motor behaviors, such as reaching and grasping and moving through the environment. David Milner and Melvyn Goodale propose that early in the evolution of animals, the major goal of visual processing was not to create a conscious perception or "picture" of the environment, but to help animals control navigation, catch prey, avoid obstacles, and detect predators—all crucial functions for the animal's survival. A large amount of physiological research influenced by this line of thinking has demonstrated close links between perceiving and taking action.

The power of the physiological approach is illustrated by the physiological mechanisms that have been proposed to explain the vast range of perceptions described in this book. Physiological principles and procedures like those described here (as well as many others that are described in individual entries), used in conjunction with psychophysical measurements, have revealed many of the physiological mechanisms that underlie perception, recognition, and action.

E. Bruce Goldstein

See also Action and Vision; Attention: Physiological; Audition; Auditory Frequency Analysis, Neural; Auditory Localization: Physiology; Auditory Processing: Central; Auditory Processing: Peripheral; Color Perception: Physiological; Cutaneous Perception: Physiology; Eye Movements: Physiological; Kinesthesia; Lateral Inhibition; Modularity; Motion Perception: Physiological; Multimodal Interactions: Neural Basis; Neural Recording; Neural Representation/Coding; Object Perception: Physiology; Olfaction; Pain: Physiological Mechanisms; Proprioception; Receptive Fields; Spatial Layout Perception: Neural; Speech Perception: Physiological; Taste; Tuning Curves; Vision; Visual Processing: Primary Visual Cortex; Visual Processing: Extrastriate Cortex

Further Readings

Barlow, H. B. (1995). The neuron in perception. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 415–434). Cambridge: MIT Press.

Goldstein, E. B. (2001). Cross-talk between psychophysics and physiology in the study of perception. In E. B. Goldstein (Ed.), *Blackwell handbook of perception* (pp. 1–23). Oxford, UK: Blackwell.

Hollins, M., Bensmaia, S. J., & Washburn, S. (2001). Vibrotactile adaptation impairs discrimination of fine, but not coarse, textures. *Somatosensory and Motor Research, 18,* 253–262.

Hubel, D. H. (1982). Exploration of the primary visual cortex, 1955–1978. *Nature, 299,* 515–524.

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology, 16,* 37–68.

Milner, A. D., & Goodale, M. A. (2006). *The visual brain in action* (2nd ed.). New York: Oxford University Press.

Mountcastle, V. B., & Powell, T. P. S. (1959). Neural mechanisms subserving cutaneous sensibility, with special reference to the role of afferent inhibition in sensory perception and discrimination. *Bulletin of the Johns Hopkins Hospital, 105,* 201–232.

Newsome, W. T., Shadlen, M. N., Zohary, E., Britten, K. H., & Movshon, J. A. (1995). Visual motion: Linking neuronal activity to psychophysical performance. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 401–414). Cambridge: MIT Press.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2008). Sparse but not "grandmother-cell" coding in the medial temporal lobe. *Trends in Cognitive Sciences, 12,* 87–91.

Rolls, E. T., & Baylis, L. L. (1994). Gustatory, olfactory, and visual convergence within the primate orbitofrontal cortex. *Journal of Neuroscience, 14,* 5437–5452.

Rolls, E. T., & Tovee, M. J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology, 73,* 713–726.

Sheinberg, D. L., & Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proceedings of the National Academy of Sciences USA, 94,* 3408–3413.

Tong, F., Nakayama, K., Vaughn, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron, 21,* 753–759.

Ungerleider, L. G., & Haxby, J. V. (1994). "What" and "where" in the human brain. *Current Opinion in Neurobiology, 4,* 157–165.

PICTORIAL DEPICTION AND PERCEPTION

Pictorial depiction, as used in this entry, refers to two-dimensional representations of environmental objects or scenes. Imagine an artist standing in front of a canvas and observing a scene that contains many different objects and variations in illumination. To produce a realistic depiction of this scene, the artist must solve the following problems: (a) The scene is three-dimensional, extending into space and with objects located at various distances, whereas the canvas is a flat, two-dimensional surface; (b) The scene has a large range of illuminations, created by sunlight and shadows, whereas the picture's range of illumination is limited by the amount of light reflected by light and dark paints; (c) The scene contains a myriad of details, far too many to depict in a small picture. In addition to these problems, the artist must decide on a point of view for the scene, what to include and what to leave out, and perhaps also whether to take some liberties with the placements and viewing angles of some of the objects in the scene.

This entry considers some of the perceptual considerations involved in solving the previous problems. In doing so, it will become clear that in the service of "realism" or "understandability," the artist must sometimes manipulate some of the properties of the image he or she is creating. This is necessary because of physical and perceptual differences between pictures and the environment. Despite these differences, artists can use perceptual principles to create realistic depictions of scenes and objects within scenes.

Depicting Objects

An important difference between perceiving an object in a scene and in a picture is that a person can move around in a scene to view the object from different angles, but in a picture the object is frozen. The viewer sees the object from the point of view selected by the artist and cannot inspect it further. The ability to view objects in scenes from different angles is important, because although many objects—especially familiar ones—appear just as one would expect them to when viewed from different angles, surprises can potentially occur. Consider, for example the metal horse in Figure 1, which appears to be a "realistic" horse from viewpoint 1 (Figure 1a), but which is revealed to be distorted when observed from viewpoints 2 and 3 (Figures 1b and c). The fact that a single view of an object provides incomplete information about an object means that when a person views objects, he or she routinely makes assumptions about the objects. When the object is in the environment, the person can check these assumptions by moving to different viewpoints, but when the

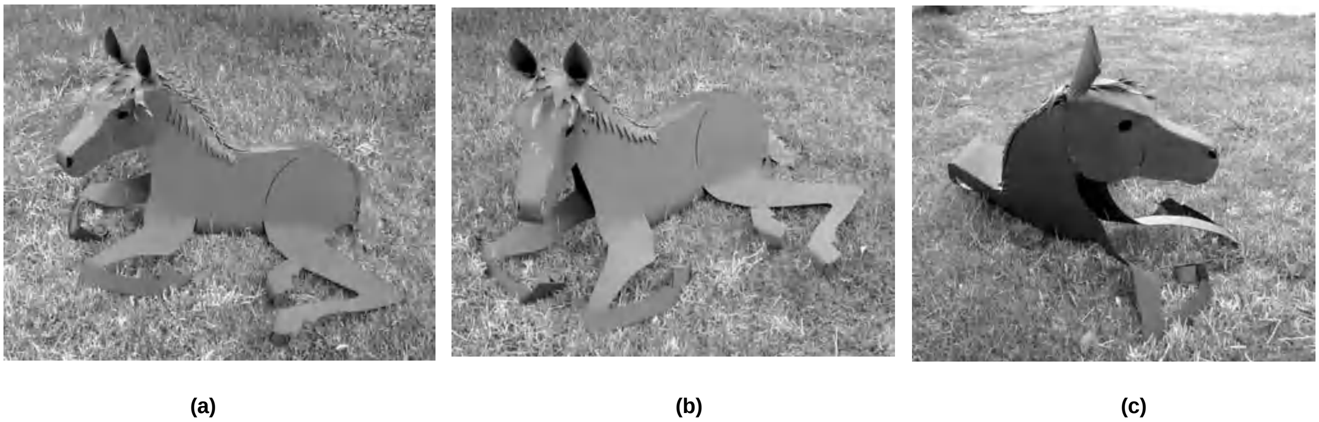


Figure 1 Three Views of a "Horse"

Note: Moving around an object can reveal its true shape.

object is in a picture, there is no way of checking one's assumptions.

Given that the information depicted in pictures is incomplete and can't be checked by changing the point of view, it is important that the artist depict individual objects as clearly as possible. One way to achieve this is to depict the object from a viewpoint that shows properties that are present at many viewing angles. For example, all three parallel edges of the book in Figure 2 are visible in the views shown in Figures 2(a) and (b). This property—three parallel edges—is called an *invariant property* of the book, because it is generally unaffected by variations in viewpoint. In contrast, the view in Figure 2(c) reveals an *accidental property* of the book, in which the three parallel edges are not visible. This situation occurs only on the rare occasions when the book is viewed so only one face is visible. Objects are difficult to identify when only accidental properties are visible, so it is important that artists avoid these viewpoints when depicting objects.

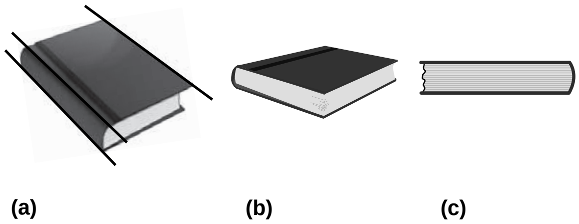


Figure 2 Invariant and Accidental Properties

Notes: One of a the book's invariant properties is the three parallel edges indicated by the darkened lines in (a). This property is called invariant because it is visible when the book is viewed from different angles, as in (a) and (b). In the view from (c) the three edges are not visible. This is an accidental property of the an book, which occurs only rarely. Clarity of depiction is enhanced by avoiding viewpoints that result in accidental properties.

The artist must also deal with the problem of depicting details. Although a building may be constructed with thousands of bricks, which a person could perceive when looking at the actual building, the artist does not have to depict each brick in order to covey the presence of a brick wall. If it is clear that the object is a building, and some bricks are shown, the viewer can fill in the other bricks based on his or her knowledge of the environment and the conventions of pictorial representation. This "filling in" of details has been called the "etcetera principle" by art historian Ernst Gombrich.

Depicting Depth and Space

Pictures of three-dimensional (3-D) scenes have been described as having a "dual reality," because the markings on the flat, two-dimensional (2-D) surface of the picture can represent 3-D objects and scenes. The challenge, therefore, is to provide information on the 2-D surface that corresponds to the information provided by 3-D objects or scenes. This problem was solved in the Renaissance by the discovery of linear perspective—a method of depiction that created a geometrical projection of a scene on the picture surface.

Figure 3 Basic Principle of Linear Perspective From Brooke Taylor, *New Principles of Linear Perspective* (1811)

Notes: Lines extend from the viewer's eye to corners of the cube. The places where these lines cross the picture plane determine the locations of the cube's representation in the picture.

Linear Perspective and the Illusion of Space

The creation of illusionistic pictorial space by Renaissance artists has been called one of the great achievements of Western history. The widespread use of linear perspective that began in the Renaissance gave artists the ability to objectively depict the geometry of space, and to gain some control over it, two outcomes very much in concordance with the Renaissance quest to understand and control nature.

The basic principle of linear perspective, illustrated in Figure 3, is to create a projection of a 3-D scene onto a 2-D surface by extending imaginary lines from the viewer's eye to the object. The places where these lines intersect a picture plane create a depiction in which each point on the picture corresponds to a point on the object. The beauty of linear perspective is that when a perspective depiction of a scene is viewed from the correct station point (with the viewer's line of sight perpendicular to the picture and the viewer positioned so the picture covers the same angle of view as did the scene viewed during the artist's creation of the picture), the picture creates an image on the observer's retina that duplicates the image created by the actual scene. Another way to enhance the perception of three-dimensionality in pictures is to include pictorial depth cues such as overlap, relative size, and atmospheric perspective.

Although linear perspective and depth cues can be used to provide information about the depth in a scene, observers viewing a picture are usually not fooled into mistaking the picture for the actual scene. The reason the observer is not fooled is because *flatness cues* provide information that makes the viewer aware of the flat surface of the picture. Examples of flatness cues are visibility of the picture's surface texture and border, and similarity of the left-eye and right-eye images the picture creates on the retina. (In contrast, 3-D scenes create slightly different images on the left and right eyes.)

Flatness cues can be eliminated by (a) eliminating surface texture which, for example, might occur in a painting by Vincent van Gogh, with its thickly applied paint; (b) eliminating visibility of the picture's borders by having the viewer look through a peephole; and (c) viewing the picture with one eye, which can also be accomplished by viewing through a peephole. Eliminating flatness cues can create the illusion that the viewer is looking, not at a picture, but at the actual scene.

Creating pictures that completely fool the viewer into mistaking the picture for a real scene may not be the primary goal of most artists, especially because people do not typically view pictures from the correct viewing point and artists do not usually intend that their pictures be viewed through a peephole. Artists have, however, created illusionistic pictures called *trompe d'oeil* (fool the eye) by creating realistic depictions of scenes that have little or no depth. For example, the American trompe d'oeil artist, William Harnett, painted pictures of objects such as musical instruments (a violin), paper (a musical score), and articles of clothing (a hat) hanging on flat surfaces, such as a door. The flat nature of the objects eliminated the need to depict deep three-dimensional space, and the result (aided by Harnett's realistic depiction of details and shadows) are pictures that look very much like actual objects hanging on a flat surface.

The Robustness of Perspective

Although it is possible to duplicate the geometrical layout of a scene when a linear perspective picture is viewed from the correct station point, pictures are most often viewed from many different distances and angles. When a picture hanging on a wall is viewed from the side, the image that the picture creates on the observer's retina becomes distorted compared to the image of the actual scene. However, viewers are usually unaware of such distortions except at extreme viewing angles. This seeming invulnerability of pictures viewed at an angle (at least at moderate angles) to distortion has been called the *robustness of perspective.*

The mechanism responsible for this resistance to distortion is unclear. One proposal is that the visual system takes the orientation of the picture's surface into account. Another proposal is that viewers don't notice distortions because of the natural variability in the shapes of most environmental objects. According to this idea, if viewing a picture at an angle makes a pictured person appear slimmer, such slimmer people actually exist, and so the distortion is not evident.

Although viewers may have only a limited awareness of distortions when viewing single pictures at an angle, awareness of this distortion becomes more obvious when the same picture is simultaneously viewed from two angles. This awareness can be demonstrated by creating two identical copies of a picture by xerography and positioning one on the left side of a page and the other on the right. By folding the page between the two images, it is possible to create a situation in which one image can be viewed at an angle while the other is viewed straight on. Comparison of the two pictures, when viewed simultaneously from different angles, reveals a large difference in the perception of the two views.

The phenomenology of this distortion aside, there is evidence that the *spatial layout* of the scene depicted in a picture (as measured by the observer's arrangement of elements of the pictured scene on a three-dimensional plane in front of the picture) is relatively unaffected by changes in viewing angle. Although the mechanism responsible for this result is unclear, this "layout constancy" is yet another example of the robustness of pictures.

Depicting Illumination

Another challenge in creating pictures that mimic properties of the environment is the depiction of illumination. The problem is that the range of illuminations in the environment can be 10,000 to 1 or greater if we compare the illumination created by direct sunlight to that of a darkened room, but the range of illumination possible in pictures is limited by the percent of light reflected from white pigment (about 95%) compared to black pigment (about 5%)—a range of about 20 to 1.

There are a number of ways to deal with this problem:

- Create contrast. The perceived brightness of a light area can be enhanced by surrounding it by a dark area. Conversely, a dark area can be made to appear darker by placing it on or adjacent to a light area.
- Include information associated with brightness in the environment. Some examples are (a) depict less detail in dim light, more in bright light; (b) depict shadows with sharp borders to indicate bright, direct light, and fuzzy borders to indicate diffuse light; (c) depict glare; (d) depict illumination as yellow, to indicate bright sunlight.
- Use other special techniques. Artists have devised other ways to simulate illumination. One of the most notable is the technique of pointillism, developed by the 19th century French impressionists, in which pictures were created using small dots of color. When viewed from certain distances, these dots create a "shimmering" effect, which adds to the perception of illumination.

Composing Pictures

In addition to using special techniques that have been used to depict objects, space, and illumination, artists can also control a picture's composition—where objects and properties such as color, lightness, and darkness are arranged in the picture. When books on art theory discuss composition, they often focus on the following perceptual effects of "good" composition: (a) achieving perceptual clarity, (b) directing the viewer's attention to various areas of the picture, and (c) creating a feeling of harmony or balance. This entry considers the first two.

Achieving Perceptual Clarity

Pictures are particularly vulnerable to ambiguity. For example, a cube depicted from a viewpoint that shows just one face may appear to be a square, or a person in the foreground of a photograph may appear to have a lamp, which is actually located in the background of the scene, growing out of his or her head. It is such pictorial "mistakes" that good composition seeks to avoid. Composition texts caution artists to avoid following or prolonging the line of the edge of one object, such as the trunk of a tree, directly into the edge of another object, such as the edge of a house. The result, called *false attachment,* can be an unwanted perceptual joining of two objects that are actually separated from one another.

Another technique that has been employed to make pictures easier to "read" is the manipulation of contrast to ensure the separation of the figure from the background. For example, the perception of an "important" object, such as a person, can be enhanced by depicting a person in dark clothing on a light background, or by depicting a person's face as illuminated and darkening the background surrounding the face. The beauty of pictures is that the artist has the luxury of being able to manipulate properties such as placement of objects and contrast. The resulting picture, although "manipulated," can look like a realistic depiction of an environmental scene.

Directing Attention

Some composition texts state that if a picture is properly composed, the viewers will never get "lost." That is, good composition can determine *where* a viewer looks, and the *order* in which the objects in the picture are viewed. Thus, a properly composed picture will draw the viewer into the picture and keep him or her from leaving it. Techniques such as depicting lines or a pictured person's gaze so they lead into the picture have been suggested as ways to control the way a viewer scans a picture. Lighting can have powerful effects as well. For example, the baroque artists Caravaggio and Georges de la Tour used *chiaroscuro*—lighting that makes forceful use of contrasting lights and darks—to direct the viewer's attention to specific areas in their paintings.

The idea that composition can direct the viewer's attention is, however, only partially supported by empirical results. Measurements of the eye movements made by observers as they view pictures support the idea that artists can determine *where* a viewer will look in a picture, because viewers tend to look at areas of pictures that are especially meaningful, such as human faces, or especially prominent features, such as high brightness or contrast. However, there is little empirical evidence to support the idea that compositional devices can control the *sequence* in which viewers look at objects in a picture. In fact, there is evidence that scan paths—the order in which a viewer looks from one glance to the next—are variable across observers, so a particular picture might elicit 10 different scan paths from 10 different observers.

The issue regarding whether pictorial composition can control the sequence of scanning is still, however, open to investigation. The answer awaits a systematic investigation of the eye movement patterns in response to paintings that have been specifically composed according to principles designed to direct eye movements.

Composition of Portraits

Recent research on how people are depicted in portraits has revealed that portraits are more often oriented so the sitter's left cheek is visible (i.e., the sitter is facing leftward in the picture, from the perspective of the person viewing the picture). This effect is a robust one that extends over centuries of Western art, with proportions of left-cheek-visible pictures being about 68% for portraits of women and 56% for portraits of men. This left-facing preference has generated speculation regarding possible mechanisms. One of the most intriguing hypotheses is that light is more likely to be depicted as coming from the left, so a sitter facing left will be turned toward the light and a sitter facing right will be turned away.

Yet another intriguing pattern in portrait painting is the placement of a person's eye so one of the eyes of frontal or three-quarter portraits is positioned so a line bisecting the picture passes through one of the eyes with an accuracy of plus or minus 5%. This eye centering occurs for portraits throughout the history of Western art, ranging from the *Mona Lisa* to the statesmen on U.S. Treasury Notes. (Check George Washington's eyes, for example, on the $1 bill, and Andrew Jackson's on the $20 bill.) Interestingly, a general principle of eye centering is mentioned only rarely in composition texts or the analytic literature of art. Because artists aren't taught to center eyes, it has been suggested that this depiction may be created unconsciously and is perhaps related to the eye's role as a channel for the visual transmission of consciousness between people.

This idea is a fitting one on which to end this discussion of pictorial depiction and perception, because one of the central facts about pictures is they are creations of one human mind that are designed to be understood by other human minds. Thus, creating a "realistic" picture involves much more than depicting what is "out there." Pictures are often the end result of a series of decisions driven both by the artist's intentions and characteristics of the intended observer's perceptual system.

E. Bruce Goldstein

See also Aesthetic Appreciation of Pictures; Contrast Enhancement at Borders; Depth Perception in Pictures/Film; Eye Movements: Behavioral; Gestalt Approach; Impossible Figures; Object Perception; Perceptual Organization: Vision; Perceptual Segregation; Spatial Layout Perception: Psychophysical; Virtual Reality: Vision; Visual Displays; Visual Scene Perception

Further Readings

Arnheim, R. (1974). *Art and visual perception* (Rev. ed.). Berkeley: University of California Press.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94,* 115–147.

Goguen, J. A. (Ed.). (1999). *Art and the brain.* Thorverton, UK: Imprint Academic.

Goldstein, E. B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 256–266.

Goldstein, E. B. (2001). Pictorial perception and art. In E. B. Goldstein (Ed.), *Blackwell handbook of perception* (pp. 344–378). Oxford, UK: Blackwell.

Gombrich, E. H. (1956). *Art and illusion.* Princeton, NJ: Princeton University Press.

Gombrich, E. H. (1982). *The image and the eye.* Ithaca, NY: Cornell University Press.

Gregory, R. L., & Gombrich, E. H. (Eds.). (1973). *Illusion in nature and art.* New York: Scribners.

Hagen, M. A. (Ed.). (1980). *The perception of pictures* (Vols. 1 & 2). New York: Academic Press.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7,* 498–503.

Hochberg, J. (1996). The perception of pictures and pictorial art. In M. F. Friedman & E. C. Carterette (Eds.), *Cognitive ecology* (pp. 151–204). San Diego, CA: Academic Press.

Ittleson, W. H. (1996). Visual perception of markings. *Psychonomic Bulletin & Review, 3,* 171–187.

Kubovy, M. (1986). *The psychology of perspective and renaissance art.* Cambridge, UK: Cambridge University Press.

Stark, L. W., & Ellis, S. R. (1981). Scan paths revisited: Cognitive models direct active looking. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 193–226). Hillside, NJ: Lawrence Erlbaum.

Tyler, C. (1998). Painters center one eye in portraits. *Nature, 392,* 877–878.

Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience, 8,* 1401–1410.

PLASTICITY

See Experience-Dependent Plasticity

POSITRON EMISSION TOMOGRAPHY

See Brain Imaging

PRIMING

In perception, *priming* denotes the mutual influence between different percepts, processes, or representations. Imagine that you can't remember where you left your coffee mug in your house, so now you search for it. If you commonly drink the coffee in the kitchen, one look into the kitchen is probably enough to find it. This quick perception exemplifies priming: Perceiving a familiar context facilitates seeing an object there. Priming is often beneficial. Yet it can also be detrimental. Imagine you search for a child in a playground. The playground is a common context for perceiving a child. Therefore, it primes perceiving children—but not just yours. In fact, the more children potentially distract you, the more harmful the playground's priming effect for finding a particular child will be. This entry discusses the types and characteristics of priming, as well as priming in perception.

Types

All priming can be considered a primitive form of memory, exploiting the principle that a preceding (priming) stimulus affects the perception of a subsequent (primed) stimulus or the behavior toward it. Different types of priming can be differentiated by considering the relationship between priming and primed stimulus. Starting with the most detailed and ending with the most abstract commonalities between priming and primed stimulus, priming can be subdivided into repetition priming, sensory priming, contextual priming, and conceptual priming.

Perceiving an identical stimulus twice creates repetition priming. Yet priming effects are observed with lesser similarity between priming and primed stimulus. In sensory priming, priming and primed stimulus are highly similar but not identical to one another. A powerful type of sensory priming uses images of the same object taken from different angles of view. Seeing a face from the front helps subsequently recognizing that face in profile. Such priming is frequently used in cinema, facilitating reorientation after a cut. Sensory priming can even cross modality boundaries: Hearing the sound that an object makes while contacting a surface can facilitate visual recognition of the surface as being fluid or solid. This type of priming is still sensory because the similarities concern representing the underlying physical characteristics: the amount of reverberation heard and seen, for example.

During contextual priming, perceiving a fitting context primes the representation of an associated object. Seeing the mug in the kitchen, as in the first example, is all about contextual priming. Spatial priming contributes to contextual priming. Spatial priming means that different locations contain different objects with different probabilities. For instance, locations in the sky frequently contain flying objects, such as birds or airplanes, but they rarely contain nonflying objects like elephants or lampposts. Thereby, spatial priming exploits ecological principles, for instance, gravity (e.g., immobile solid objects are usually found on solid surfaces).

Note that many sorts of priming reflect the repeated "use" of representations for similar purposes, rather than the similarities of the representations themselves. In sensorimotor priming, perceiving an object facilitates an appropriate motor action. The effect is striking during observation of conspecifics (e.g., yawning is contagious). Sensory and sensorimotor priming are fundamentally related. Consider recognizing a spoon. Is it the visual image that makes you recognize it is a spoon? Or do you recognize the spoon by the motor acts that it affords, such as its usability for stirring? James Gibson advocated the view that motor affordances—the specific actions that a particular object affords for a human observer—are the basis of perception. Taking an ecological perspective this makes sense. Perception is a means to an end: It provides sensory representations for fitting behavior.

Finally, priming can be more abstract or arbitrary than in the previous examples, with only idiosyncratic relations between priming and primed stimulus. Conceptual priming does not have to draw on any underlying ecological or physical regularities. Instead, it can completely depend on an observer's idiosyncratic personal experience. Have you ever recognized an old classmate you haven't seen since long ago by his voice? Here, an auditory representation helps you notice visual similarities between the perceptual representation and the memory trace. However, this voice and this particular image are not associated by means of any ecological or physical regularity. Instead, conceptual priming hinges solely on your idiosyncratic experience with the particular instance or token bearing the association.

Sensorimotor priming can be also abstract. If asking an observer for an arbitrary response to a specific shape, a sufficiently similar shape can also prime this arbitrary response. During so-called overgeneralization of applying a category to objects, for instance, a child might give the arbitrary vocal response “dog” to any object sufficiently similar to a dog, for instance, one that has fur, four legs, and moves, like a cat or a horse.

Characteristics

Concerning priming’s temporal characteristics, an interval of some milliseconds between priming and primed stimulus suffices for a priming effect. However, the impact of priming is seen for a much longer duration. Priming effects can last up to several months after the priming stimulus. Such long-lasting priming effects partly reflect unconscious memories because they do not require a conscious recollection of the priming stimulus.

Nonetheless, long-lasting priming effects seem to require a conscious perception of the priming stimulus. This is indicated by the observation that priming effects of invisible stimuli are only short lived. The priming effect of a word presented below the level of conscious perception, for instance, declines to zero within 300 to 700 milliseconds after the word’s beginning.

Concerning priming’s spatial characteristics, the extent of priming across space scales with the visibility of the stimuli: A visible image can have a priming effect on a subsequent similar image, even if priming and the primed image are shifted relative to one another by about 7° of visual angle. By contrast, an invisible image’s priming effect is restricted to spatial shifts of about 4° or less.

Both consciously and unconsciously registered stimuli exert a priming effect. However, temporal and spatial properties of the priming effect suggest that conscious stimuli bring about a more encompassing (i.e., spatially and temporally generalized) priming effect.

Priming in Perception

In perception, priming as an effect is brought about by two mechanisms, repetition and preparation. An example for priming by repetition is the way that neurons (nerve cells) in the brain represent a visual stimulus. Many visual neurons just represent one particular feature (e.g., a specific color). Such a neuron only fires if it is triggered by a fitting input. Importantly, however, each neuron has a voltage threshold regulating when it fires: A neuron only fires if its input voltage is large enough for exceeding its threshold. Now, stimulating a neuron once can decrease the threshold of that neuron. Such repetition priming in neurons can account for sensory priming.

For preparation to take effect, an association between two features is necessary and sufficient. The features do not have to resemble one another. Seeing one of a multitude of associated features, an observer can prepare processing of the other expected or predicted features. Imagine you visit a zoo. When searching for a mouse in a cage, you probably scan the ground, maybe focusing on holes (i.e., good hiding places). When searching for a hummingbird, by contrast, you probably look in the air, maybe focusing on blossoms on trees (i.e., good feeding places). In both examples, associations connecting animals with specific behavior and locations can be used for preparing the quick perception of the searched animal. Generally, preparation can also account for conceptual priming, that is, the priming effect based on the more idiosyncratic expectations of the observer.

Preparation priming comes in two varieties. On one hand, it can reflect the observer’s conscious expectation of a feature *y* after perceiving an associated feature *x*. In this case, preparation priming is a way that an observer willingly or intentionally improves her or his perception. On the other hand, several instances of preparation priming probably do not critically hinge upon conscious expectations. This is clear from priming effects occurring without any conscious recollection of the underlying association. A good example is amnesia. During amnesia, a patient cannot explicitly recall a seen stimulus. If an amnesic is shown the word *traffic,* for example, he or she cannot remember if asked to recall whether he or she has previously seen the word. Yet when prompted to complete the letters “TRA” into a word, the same amnesic will use the associated letters, that is, answer with the seen word *traffic.*

Priming can precede the conscious perception of a stimulus. Therefore, priming is a mechanism suited for improving subsequent perception. The fact that priming draws on mechanisms ranging from feature-driven repetition to willingly controlled preparation of processing a feature shows that priming could even be an exceptionally general principle for boosting perception.

Ulrich Ansorge

See also Context Effects in Perception; Direct Perception; Neural Representation/Coding; Multimodal Interactions: Spatial Perception in Touch and Vision; Speed of Processing in Sensory Systems; Unconscious Processes; Visual Scene Perception; Visual Scene Statistics

Further Readings

Ansorge, U., Neumann, O., Becker, S., Kälberer, H., & Cruse, H. (2007). Sensorimotor supremacy: Investigating conscious and unconscious vision by masked priming. *Advances in Cognitive Psychology, 3,* 257–274.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Science, 11,* 280–289.

Schacter, D. L., Wig, G. S., & Stevens, W. D. (2007). Reductions in cortical activity during priming. *Current Opinion in Neurobiology, 17,* 171–176.

PRISM ADAPTATION

When people first look through prisms that displace the visual world left or right, they experience errors in their behavior. For example, they reach to the side of their coffee cup and may bump into a door frame when walking. With continued exposure to the prisms, however, these errors disappear. The person has *adapted* to the prisms. When the prisms are removed, the person may initially make similar errors, but now in the opposite direction. These *aftereffects* demonstrate that a change has occurred in the person's perceptual-motor system (i.e., the organization of senses and limbs), a change that persists until the person relearns the normal manner of interacting with the world. Once an amusing curiosity good for classroom demonstrations, prism adaptation is proving to be a valuable tool for investigating adaptive perceptual-motor performance and the rehabilitation of brain-damaged patients. This entry describes some of the mechanisms that have been proposed to account for prism adaptation and a practical application of prism adaptation.

Research has shown that prism adaptation involves four kinds of adaptive processes. (1) The person may consciously guide the visible hand toward the coffee cup, in which case, aftereffects will *not* occur. (2) The person may unconsciously adjust head posture so that the head is turned in the direction of the displacement, in which case the coffee cup appears and *is* straight ahead, performance error is reduced, and head posture aftereffects will occur. (3) The person may learn to correct for the prisms in the same way that everyday reaching errors are corrected: Aftereffects of such perceptual-motor learning can occur, especially if the person does not recognize that the prisms have been removed. (4) The person may correct the underlying spatial mappings that are disrupted by the prisms. This last kind of adaptive change is the unique feature of prism adaptation.

Current research suggests that our perceptual-motor system is composed of multiple sensorimotor systems, each of which can operate independently of the others, but which can be *coordinated* to perform a more complex task. For example, the eye-hand perceptual-motor system includes the visual eye-in-head sensorimotor system, which is exercised autonomously when we sit quietly reading, and the proprioceptive hand-to-head sensorimotor system, which operates autonomously when we struggle to find the alarm clock in a dark room. Coordinating the eyes and hand to pick up the visible morning coffee cup poses a problem because visual and proprioceptive systems do not have the same spatial maps.

When we stand up, our eyes and limbs change position relative to the coffee cup, and we have to change how we look at and reach for the coffee cup on the table. Similarly, our eyes and limb have different origin positions: The eyes are located in the head and the limb is attached at the shoulder. The seen coffee cup is located relative to our head, whereas the cup in hand is located relative to our shoulder. This difference in cup location relative to the head and shoulder must be taken into account when we reach for the seen cup. The position of the seen cup must be adjusted for the distance between the head and shoulder if the hand is to successfully pick up the cup: That is, position in visual space must be *transformed* into position in proprioceptive space or vice versa if we want to see if there is coffee in the cup in hand.

These spatial transformations are already in place for the normal adult, and sensorimotor systems are spatially aligned. Prisms, however, change the needed visual transformation, and prism adaptation requires realignment. Depending upon the conditions of exposure that determine which system is guiding another, spatial realignment will occur in different sensorimotor systems. For example, in normal visually guided reaching, the proprioceptive transform is realigned with the visual system, and aftereffects occur in unsighted reaching. On the other hand, for proprioceptive guided looking at the cup in hand, the visual system is realigned with the proprioceptive system and visual aftereffects occur. Under most conditions, some of both kinds of realignment occur and aftereffects *add* to equal to the total aftereffect for the complete eye-hand coordination loop.

Recently, prism adaptation has been found to have application in the treatment of unilateral neglect. This affliction sometimes occurs with damage to the right cortical hemisphere in the region of posterior parietal lobe, near the interface with the parietal temporal lobe. The hallmark of neglect patients is the lack of awareness of the left half of space: they may eat only from the right side of their plate; they may dress only the right side of their body; if ambulatory, they may run into the left side of the door or wheel their chair only with their right hand. And, neglect patients vehemently deny that they have a problem. “What do you mean there is more on the left? I see everything!” Rehabilitation of such patients has been problematic. However, a single 15 minute rightward prism adaptation session can markedly reduce clinical symptoms and the amelioration may persist for hours, days, or even weeks and may even increase without further treatment!

These results suggest that the brain site of spatial realignment is subcortical, likely cerebellar, because patients with cortical damage can adapt to prisms. Neglect likely arises from damage to the cortical brain structures responsible for strategic perceptual-motor control. One hypothesis is that neglect patients have, in part, forgotten the perceptual-motor programs that allow interaction with the left half of the world. Prism adaptation forces a change in the intact alignment transforms to bring about interaction with the left half of the world again; and patients begin to remember the forgotten programming. Having once remembered, even partially, the patient can then remember more, even without the prisms.

Gordon M. Redding

See also Consciousness: Disorders; Perceptual Learning; Perceptual-Motor Integration; Vision: Developmental Disorders

Further Readings

Bedford, F. (1993). Perceptual and cognitive spatial learning. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 517–530.

Danckert, J., & Ferber, S. (2006). Revisiting unilateral neglect. *Neuropsychologia, 44,* 987–1006.

Newport, R., & Jackson, S. R. (2006). Posterior parietal cortex and the dissociable components of prism adaptation. *Neuropsychologia, 44*(13), 2757–2765.

Redding, G. M., Rossetti, Y., & Wallace, B. (2005). Applications of prism adaptation: A tutorial in theory and method. *Neuroscience and Biobehavioral Reviews, 29*(3), 431–444.

Redding, G. M., & Wallace, B. (1997). *Adaptive spatial alignment.* Mahwah, NJ: Erlbaum.

Redding, G. M., & Wallace, B. (2006a). Generalization of prism adaptation. *Journal of Experimental Psychology: Human Perception and Performance, 32*(4), 1006–1022.

Redding, G. M., & Wallace, B. (2006b). Prism adaptation and unilateral neglect: Review and analysis. *Neuropsychologia, 44*(1), 1–20.

Welch, R. B. (1978). *Perceptual modification: Adapting to altered sensory environments.* New York: Academic Press.

PRIVATE NATURE OF PERCEPTUAL EXPERIENCE

Two people look at an apple. What is their experience of the apple’s color and how do their two experiences compare? If they both say the apple looks “red,” does this mean their perceptions are the same? This entry presents evidence supporting the idea that it is likely that people experience the same stimuli differently, and also considers the idea that because the essence of our experience is essentially private and can’t be shared with others, it is difficult or impossible to determine when people’s perceptions are the same. Related to the private nature of human perception is the issue of whether it is possible to know what animals experience.

What Does Labeling Experience Indicate About Experience?

What does it mean if two people say "The apple looks red," or, using another descriptive technique, "It is the same color as blood"? Does this provide any information about what the people are experiencing? To a certain extent it does. Their description indicates that they label what they are experiencing as "red," and that their color-experience caused by the apple is similar to their color-experience caused by blood. But labeling or comparing to another experience does not capture the true essence of experience. The essence of our experience—what it is like to experience red, feel a pinprick, or hear the sound of a violin—is called *qualia*. The experience of red, looked at in this way, can be described as having a red *quale*. Thus, our question becomes, "Is one person's red quale the same as another person's red quale?" It is clear that their response of "red" to the apple is no more than a label for each person's "red-quale" experience. It could be that one person's red quale is a different shade of red or perhaps a completely different color altogether, compared to the other person's red-quale.

The idea that two people's qualia might be different may appear counterintuitive because of "shared experiences," such as one person saying "the sky is a beautiful shade of blue" and another person agreeing. Although this agreement may imply a shared experience of "blueness" or "beauty," there are a number of reasons why it is reasonable to expect that different people could experience the same stimuli differently. This conclusion derives from both physiological and behavioral considerations.

Physiological Reasons to Expect Differences in Perception

There are many examples, which extend across senses, of situations in which physiological differences between people might cause differences in perception. The basic idea is that just as there are physiological differences between people's appearance, physical strength, and susceptibility to certain genetically based conditions, there are also physiological differences in the mechanisms responsible for perception.

Physiological Differences in Color Vision

Color perception is due to the action of three types of photosensitive chemicals, called visual pigments, that are each contained in separate visual receptors. These receptors have been characterized as short-wavelength, medium-wavelength, and long-wavelength receptors to indicate the region of the visible spectrum of light energy that causes maximum activation of each type of receptor. Perception of a particular color has been linked to the pattern of activity that light creates in the three types of receptors. For example, light that a person with normal color vision would describe as "green" would cause maximum activation in the medium-wavelength receptor and less activation in the short-wavelength and long-wavelength receptors.

The idea that different people might perceive the same wavelength of light differently follows from the finding that although people with "normal" color vision may have the same three pigments, there are large differences in the relative concentrations of these pigments and the way the receptors for each pigment are distributed across the retina. These differences could cause two people to have different patterns of cone-receptor firing to the same light, and therefore perhaps different perceptions of color.

Besides the possibility that differences in physiological makeup could cause differences in perception, there is another, more basic, physiological reason to entertain the possibility that people's perceptual experience could be different—perception can be considered a construction of the nervous system.

Perception as a Construction of the Nervous System

This entry has noted that the perception of color is determined by the patterns of neural firing in the three receptors, with blue, for example, being associated with a particular ratio of responses of the three receptors, and red, with a different ratio. But is there anything intrinsically "blue" about the "blue ratio," or "red" about the "red ratio" of receptor response? Isaac Newton realized, long before the physiological basis of color vision was known, that the light that results in the perception of color does not itself contain any color. What this means is that different wavelengths of light have no color in themselves, but that color is created by the way these wavelengths stimulate the nervous system. Color, therefore, may be related to different wavelengths of light, but there is no a priori reason to expect a particular wavelength to result in blue, another to result in red, and so on. The "blueness" or "redness" (or to be more precise, the blue or red qualia) are *created* by the nervous system, and taking this idea one step further, it is not unreasonable to propose that the nature of the qualia created in response to different wavelengths of light is arbitrary.

A similar situation exists for the sense of taste. Sucrose tastes "sweet" and quinine "bitter," and the differences between these experiences have been associated with differences in the molecular makeup of sucrose and quinine. But these molecules are no more than arrangements of atoms, such as carbon, hydrogen, sodium, and chlorine. The question then becomes, where is the bitterness in quinine's molecules or the sweetness in sucrose's molecules? The answer is that the molecules aren't bitter or sweet, but the way they stimulate the receptors on the tongue results in *experiences* of bitterness or sweetness. Bitterness or sweetness are, like colors, creations of the nervous system.

An analogous situation exists for smell. We perceive some substances as "sweet smelling" and others as "rancid," but where is the "sweetness" or "rancidity" in the molecular structure of the substances that enter the nose? Again, the answer is that these perceptions are not in the molecular structures. They are created by the action of the molecular structures on the nervous system.

Because the "blue" in a light stimulus, "saltiness" in a taste stimulus, or "sweetness" in an olfactory stimulus are contained not in the stimuli themselves, but are creations of the nervous system, there is no a priori reason to assume that each individual's nervous system will create exactly the same experience.

Behavioral Reasons to Expect Differences in Perception

There is evidence, based on measurement of behavioral responses to color and taste, that supports the idea of differences in perception between people.

Behavioral Differences in Color Vision

Behavioral evidence for differences in normal color vision is provided by color matching experiments in which a person is presented with two fields, a test field that contains a single wavelength of light, and a comparison field that consists of three different wavelengths of light mixed together. The observer's task is to adjust the amounts of the three wavelengths in the comparison field to create a color that matches the color of the test field. Results of these experiments indicate that people with normal color vision often chose different proportions of the three wavelengths. Also, two fields that are adjusted so they appear identical to one person can appear slightly different to another person.

Another test of color vision requires that observers pick "unique hues"—wavelengths in the spectrum that result in the best or purest blue, green, and yellow. Unique blue is at a wavelength of about 470 nanometers (nm), green at 510 nm, and yellow at 580 nm. However, comparisons across groups of people reveal a wide range of wavelengths for each unique hue. For example, although one study reports an average wavelength of the unique hue for green to be 515 nm, values for a group of 50 subjects ranged from 490 to 535 nm.

Behavioral Differences in Taste

Two people eat the same food. One likes it, the other doesn't. What causes this difference? On possibility is that the foods taste "sour," but one person likes the sour taste and the other person doesn't. In this case, the taste *experience* might be the same, but taste *preferences* differ. Another possibility is that the people are experiencing different tastes. This could be caused by the fact that just as concentrations of visual pigments can differ in people's visual receptors, the relative proportions of different types of taste buds differ on people's tongues. In fact, there are people with an abundance of taste buds, called supertasters, who are more sensitive to bitter than people with fewer receptors.

The conclusion from these examples from both vision and taste is that there are both physiological and behavioral reasons to expect that *differences* in perception might occur. But what about the possibility of demonstrating that two people might have the *same* experience in response to a specific stimulus? As we will now see, this problem is more difficult than demonstrating differences, and may, in fact, not be solvable because perception is basically a private experience.

Perception as a Private Experience

The evidence previously described provides reasons to believe that two people can have different experiences to physically identical stimuli. But what about demonstrating similarity? How can we determine if two people's qualia are *similar?* It has already been pointed out that applying similar labels to colors and agreeing that two stimuli look the same do not provide information about a person's qualia.

Another example illustrating the difficulty in determining whether experiences are similar is provided by the perception of pain. If two people are asked to rate a painful stimulus on a scale of 1 (not painful) to 10 (extremely intense pain), two people might assign different numbers to the same stimulus. But if one person rates their pain as a "4" and the other as a "10," how can we know whether their experience of pain intensity is different or if they have different tolerances for different levels of pain? The difficulty—or some might say the impossibility—of answering questions like this and analogous questions about color, taste, and other perceptions, reflects the fact that the essence of people's experience is private and inaccessible to another person. Thus, although we may assume that two people who call an apple red or rate pain as "4" are perceiving the same thing, this is no more than an assumption.

The Nature of Animal Experience

The idea that animals may experience the world differently than humans is generally accepted, both because of the differences in physiological makeup of humans and animals, and also because of the behavioral evidence that some animals are more sensitive to specific stimuli than humans (dogs and dolphins hear high frequencies that are inaudible to humans; bees sense short-wavelength ultraviolet light that humans can't see).

But animal examples raise some interesting questions. Consider, for example, the honeybee, which is sensitive to short wavelengths that are out of the range of human vision. Physiologically, this occurs because the honeybee has a visual pigment that absorbs light at very short wavelengths. But knowing that the honey bee can *sense* a 350-nm light doesn't tell us how the honeybee *experiences* the 350-nm light. The honeybee can't cloud the issue by saying "blue," but even so we might be tempted to assume that because humans perceive light at the short-wavelength end of the spectrum as blue, that honeybees do as well. Although it might be tempting to make this assumption, there is really no way of knowing what the honeybee is seeing. Because there is no color in the wavelengths, the bee's nervous system creates its experience of color. For all we know, the honeybee's experience of color at short wavelengths is quite different from ours, and may also be different for wavelengths in the middle of the spectrum that humans and honeybees can both see.

One of the most intriguing examples of nonhuman experience is that of the bat. Many species of bats sense by sending out high-frequency sonar waves that are far beyond the range of human hearing—and then collecting information about their environment by sensing how these waves are reflected from objects. A wave that returns quickly indicates a closer object than a wave that returns with a longer delay. Other qualities of these reflections enable bats to determine the sizes and locations of objects, and particularly importantly, for the bat's diet, to sense small bugs that can be detected even in complete darkness.

But what is the bat experiencing? The philosopher Thomas Nagel posed this question in a famous paper titled "What Is It Like to Be a Bat?" and answered that because the bat's qualia is inaccessible to us, the answer is impossible to determine. It is also interesting to consider that although the bat is using reflected waves to achieve what humans equate with "seeing," the waves themselves are more similar to the pressure changes that result in the human experience of "hearing." So does the array of reflections received by the bat conjure up perceptions something like a visual image, or does it result in a symphony of sounds, or perhaps something different from either of these familiar human perceptions?

Because activity in the bat's nervous system creates its experience, there is no way we can access the essence of the bat's experience. Animal experience, even more so than human experience, is something we can speculate about, but cannot be certain of.

E. Bruce Goldstein

See also Color Perception; Computer Consciousness; Consciousness; Electroreception; Mary the Color Scientist; Pain: Assessment and Measurement; Pain: Cognitive and Contextual Influences; Qualia

Further Readings

Abramov, I., & Gordon, J. (2005). Seeing unique hues. *Journal of the Optical Society of America, 22A,* 2143–2153.

Bartoshuk, L. M. (1980, September). Separate worlds of taste. *Psychology Today, 243,* 48–56.

Bloch, N. (1997). *The nature of consciousness: Philosophical debates.* Cambridge: MIT Press.

Brainard, D. H., Williams, D. R., & Hofer, H. (2008). Trichromatic reconstruction from the interleaved cone mosaic: Bayesian model and the color appearance of small spots. *Journal of Vision, 8*(5):15, 1–24.

Goldstein, E. B. (in press). Perception. In N. P. Azari (Ed.), *Encyclopedia of sciences and religions.* New York: Springer.

Jackson, F. (1982). Eipphenomenal qualia. *Philosophical Quarterly, 32,* 127–136.

Nagel, T. (1974). What is it like to be a bat? *Philosophical Review, 83,* 435–450.

Ramachandran, V. S., & Blakeslee, S. (1998). Do Martians see red? In *Phantoms in the brain* (chap. 12, pp. 227–257). New York: William Morrow.

PROPRIOCEPTION

The term *proprioception* refers to the representation of the ongoing configuration of the body. Charles Sherrington originally coined the term to refer to the internal representation of the body based on signals from mechanoreceptors, including spindle receptors and Golgi tendon organs in muscles, receptors in the skin, and in internal tissues and organs. He also included under proprioceptive mechanisms the vestibular system, which provides information on the angular and linear accelerations (including gravity) acting on the body and on head orientation relative to gravity. When we move about and stand on earth, our locomotion and posture are controlled against the acceleration of gravity, which is pulling our body toward the surface of the earth. Muscle activity is necessary to maintain a posture or stance against gravity, regardless of whether we are seated, standing, or walking. Awake behaving animals have postural tonus and a complex set of reflexes that allow them without conscious thought to maintain a body orientation with respect to gravity. Sherrington, in his classic book, *The Integrative Action of the Nervous System,* elucidated many of these reflexes and their underlying neural circuitry.

The term *kinesthesia* is sometimes equated with proprioception. Although the domains to which they refer overlap, it is best to keep them separate. Kinesthesia emphasizes the perception of limb motion, both active and passive, as well as the sense of effort associated with voluntary movements and the sense of fatigue. Proprioception concerns ongoing relative body configuration and body orientation to gravity, both under static conditions and during voluntary movements. There is considerable overlap in the sensory receptors and central nervous system mechanisms that underlie kinesthesia and proprioception. This entry describes the receptors contributing to proprioception, muscle spindles and adaptation to unusual force environments, visual-proprioceptive interactions, and the calibration of body dimensions.

Receptors Contributing to Proprioception

Our joints and muscles have a variety of sensory receptors that convey information to the central nervous system and that participate in various reflexes. For example, when a skeletal muscle's tendon is tapped, the muscle reflexively contracts. This tendon jerk response is used by neurologists to test the integrity of proprioceptive pathways. In the 1950s and 1960s, physiologists thought that proprioception depended virtually entirely on receptors associated with the joints of the body. These receptors were thought to act as potentiometers giving a veridical readout of the angular position of the joints. It was postulated that by combining the signals from the different joints, it would be possible for the brain to reconstruct the relative configuration of the entire body.

In the 1970s, however, a number of observations were made that called this straightforward hypothesis into question. At that time, joint replacement surgery became possible. It was soon found that after total replacement of the finger, hip, or knee joints, a process that destroyed all the receptors within the joints, accurate position sense of the involved body segments was retained. A seminal observation in 1972 helped explain why this would be the case. Peter Matthews and his colleagues at Oxford University in England made a startling discovery. They found that muscle spindle receptors within a striated muscle influence the apparent orientation of the limb controlled by the muscle. A striated muscle has two kinds of muscle fibers, extrafusal muscle fibers that do the work of contraction and are innervated by the alpha motor neurons of the spinal cord, and intrafusal muscle spindle fibers. The spindle fibers are complicated and interesting structures. They have two types of sensory endings and an independent motor innervation by the gamma motor neurons of the spinal cord. The gamma motoneuronal innervation allows the sensitivity of the receptors within the muscle spindles to be modulated (see Kinesthesia, Figure 1). When a skeletal muscle is mechanically vibrated with a physiotherapy stimulator, its muscle spindle receptors will be activated and will cause the vibrated muscle to contract and move the limb it controls. This phenomenon is known as a tonic vibration reflex and occurs because the spindle fibers monsynaptically activate the alpha primary motoneurons of the vibrated muscle. Matthews and his colleagues discovered that if they resisted the action of a tonic vibration reflex in the biceps muscle by physically restraining the forearm, the unseen stationary forearm would seem to move into extension (see Figure 1). This illusion of change in position and motion demonstrated conclusively that muscle spindle receptors affect the conscious appreciation of limb position.

The way this mechanism normally works is that the alpha and gamma motor neurons are coactivated and the nervous system computes limb position based on sensory feedback occurring during and after a movement. The brain relates patterns of alpha and gamma motoneuronal activity and relates this to the patterns of discharge from muscle spindle primary and secondary receptors, which relay velocity and length dependent responses, respectively. Coupled with information from Golgi tendon organs about developed muscle tension levels, the ongoing configuration of the body can be computed.

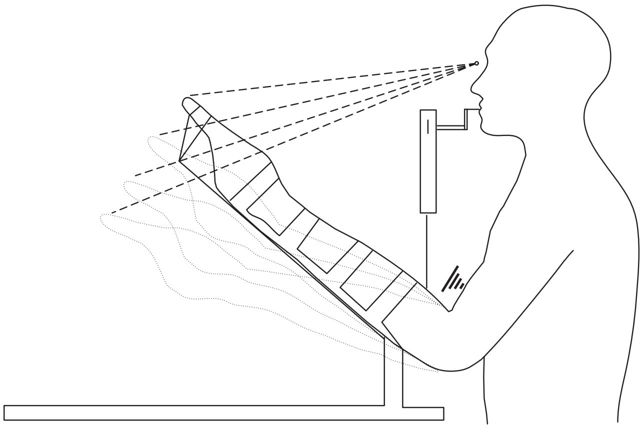


Figure 1 Illusory Movement of the Forearm Due to the Tonic Vibration Reflex

Notes: This figure illustrates the illusory motion of the forearm experienced in the dark when the arm is restrained in position and a tonic vibration reflex is induced by mechanically stimulating (dashed triangle) the biceps muscle. The dashed lines show that the person's eyes will follow the unseen, changing apparent position of the hand.

Muscle Spindles and Adaptation to Unusual Force Environments

Muscle spindle sensitivity or gain is affected by background force level and this explains why a waking astronaut may initially not sense his or her limb position. It has been found in parabolic flight experiments, in which an aircraft is flown in a roller-coaster pattern to generate alternating periods of weightlessness and of double earth gravity acceleration levels, that tonic vibration reflexes are greatly suppressed in weightlessness and heightened in intensity in high-acceleration backgrounds compared to normal earth gravity. Our bodies do not need muscle tone to maintain a stable configuration in weightlessness, but much greater than normal tone is necessary when we are exposed to increased acceleration levels. It turns out that spindle gain, which affects tonus, is automatically adjusted according to background force level. The weightless astronaut does not have touch or pressure cues on his or her body to signal orientation and the reduced or absent spindle signals are inadequate to provide information about limb configuration. This is why an astronaut on awakening in space flight may have no awareness of limb position without looking or tensing the limb muscles.

When we move one of our limbs, we sense its movement path. Muscle spindles contribute to the representation of this spatial path and to its control. An easy way to see this is in experiments in which individuals are exposed to constant velocity rotation in a fully enclosed chamber. In this circumstance, they feel completely stationary after they have been at constant velocity for a minute or so. But, when they make a reaching movement, there is an unusual force generated on their arm—a Coriolis force, as illustrated in Figure 2. Coriolis forces arise when objects move within a rotating environment and cause deviation of the objects' paths opposite to the direction of rotation. Thus, in a rotating room, the Coriolis force will displace the reaching arm laterally, opposite the direction of room rotation.

When first reaching after the onset of rotation, the individual will feel there is a force deviating the arm; that the arm did not do what was intended. However, when 8 to 10 additional reaches are made, even without visual feedback, the reaches will become straighter and more accurate. After 20 to 30 reaches, baseline accuracy per rotation will be regained, and the arm movement will no longer feel unusual in any way. The Coriolis force generated during the reaching movement is no longer sensed, even though it is still present. When the room is stopped and the person makes a movement while stationary, the movement path will again be deviated, but this time in mirror image fashion to that first reach during rotation. This movement aftereffect is the result of a central nervous system compensatory adaptation that is achieved during the rotation period. In the post-rotation period, the same central compensation that was generated during rotation is still being employed. When the post-rotation reaching error is made, it feels as if there is a force that is deviating the arm. In this case, it is the person's own, no longer appropriate, central compensation that is interpreted as an external force affecting the control of the body. It is the muscle spindle receptors that are involved in this recalibration of movement trajectory because when the movement is perturbed, they are activated and signal the change in limb trajectory relative to that intended. The touch receptors of the finger also contribute to the recalibration of the endpoint of the movement because they signal where the finger is relative to the shoulder.

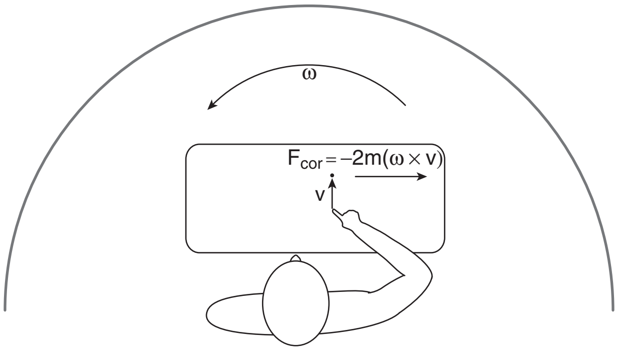


Figure 2 The Coriolis Force

Notes: This figure shows that when a reaching movement is made in a room rotating at constant velocity, ω, a Coriolis force, F_{cor}, is generated on the reaching arm. F_{cor} depends on the mass of the arm, m, the forward velocity of the arm, v, and acts opposite the direction of room rotation.

Visual-Proprioceptive Interactions

Vision and proprioception reciprocally influence one another. A simple way to appreciate this is to look at your hand through a pair of prism spectacles that laterally displace its image. In this situation, your hand will be felt to be in its seen location; in other words, vision dominates. However, the interaction can be reciprocal as the following example shows. Consider an individual seated in the dark with his or her elbow resting on a table surface and forming a 90° vertical angle and with the forearm physically restrained in position. If the biceps muscle is stimulated with a physiotherapy vibrator to elicit a tonic vibration reflex, the biceps will contract but will be unable to move the restrained forearm. In this circumstance, the forearm will feel as if it has moved into extension by as much as 60 to 70° (see Figure 1). Interestingly, if a small target light is attached to the hand, and the biceps is again vibrated, both motion of the hand and of the target light will be experienced. The apparent displacement of the hand will be about 50% smaller than when the target light is absent, and the target light will be seen to move in the same direction as the felt hand movement but by not as great a distance. In this circumstance, both the representations of hand and arm position and of visual direction are being simultaneously affected, indicating a genuine multimodal interaction. By attaching a target light to each of the hands (in an otherwise dark chamber) and creating with vibration an illusory motion of one or both hands, a change in the apparent separation of the two retinally stable targets can be induced. Thus, visual direction is remapped by proprioceptive information. The important point here is that proprioception may be the way in which visual direction is calibrated and updated in the first place. Such calibration and updating is necessary because the dimensions of the body change over time.

Calibration of Body Dimensions

An issue that arises is: How do we know the dimensions of our body, the length of our arms, the girth of our waist, and other body dimensions? One possible answer is through sight of the body but, as just mentioned, vision also has to be calibrated and proprioception seems to be involved in that process. One clue has come from studies making use of tonic vibration reflexes. If an individual is grasping his or her nose with the fingers of one hand and a tonic vibration reflex (TVR) is elicited in the biceps muscle of the grasping arm, illusory extension motion of the forearm will be induced and the nose will be perceived as getting longer, like Pinocchio's (see Figure 3). If the triceps muscle is vibrated, he or she will experience flexion of the forearm and feel his or her nose being displaced into the inside of the head. If the person is seated arms akimbo with hands holding the waist, bilateral vibration of the biceps muscles will lead to illusory extension of the forearms and an apparent growth in girth of the waist; bilateral triceps vibration will lead to apparent flexion of the forearms and an apparent wasplike waist.

The key point is that both contact of the hands with the body and proprioceptive information about the ongoing configuration of the limbs influence the perceived dimensions of the body. If the proprioceptive information is biased or inaccurate, the perceptual representation of the body is similarly affected. The question, of course, arises: How is the proprioceptive representation of hand position calibrated? We saw earlier that one way is through contact with surfaces. When one touches down on a surface with the index finger of the hand, the reaction contact force vector points to the shoulder, thus specifying the location of the finger. The important message here is that proprioceptive signals are critically important for the control of posture and movement under normal circumstances and for recalibrating movement and postural control when the physical force environment is modified, as during exposure to different force backgrounds in parabolic flight maneuvers or spaceflight. Proprioceptive signals are also crucially important for calibrating visual direction and the apparent dimensions of the body itself.

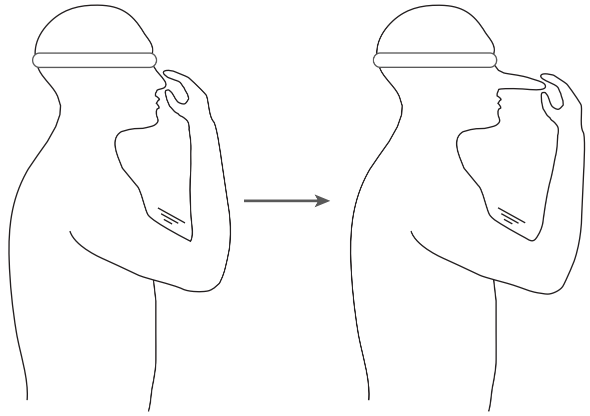


Figure 3 Pinocchio Illusion

Notes: This figure shows the Pinocchio illusion induced when a person's biceps muscle is vibrated to elicit a TVR while he or she is grasping the nose with the fingers of the vibrated arm.

James R. Lackner

See also Action and Vision; Body Perception; Kinesthesia; Multimodal Interactions: Visual–Haptic, Perceptual-Motor Integration

Further Readings

DiZio, P., Lackner, J. R., & Lathan, C. E. (1993). The role of brachial muscle spindle signals in assignment of visual direction. *Journal of Neurophysiology, 70*(4), 1578–1584.

Edin, B. B., & Johansson, N. (1995). Skin strain patterns provide kinaesthetic information to the human central nervous system. *Journal of Physiology, 487*(Part 1), 243–251.

Gandevia, S. C., Rowell, S. B., Rowell, L. E., & Shepard, J. T. (1996). Kinesthesia: Roles for afferent signals and motor commands. In *Handbook of physiology. Exercise: Regulation and integration of multiple systems. Neural control of movement* (Sect. 12, Pt. 1., pp. 128–172). Bethesda, MD: American Physiology Society.

Goodwin, G. M., McCloskey, D. I., & Matthews, P. B. (1972). The contribution of muscle afferents to kinaesthesia shown by vibration-induced illusions of movement and by the effects of paralysing joint afferents. *Brain, 95*(4), 705–748.

Knoblich, G., Thornton, I. M., Grosjean, M., & Shiffar, M. (2006). *Human body perception from the inside out.* Oxford, UK: Oxford University Press.

Lackner, J. R. (1988). Some proprioceptive influences on the perceptual representation of body shape and orientation. *Brain, 111,* 281–297.

Lackner, J. R., & DiZio, P. (1994). Rapid adaptation to Coriolis force perturbations of arm trajectory. *Journal of Neurophysiology, 72*(1), 299–313.

Lackner J. R, & DiZio, P. (2005). Motor control and learning in altered dynamic environments. *Current Opinion in Neurobiology, 15*(6), 653–659.

Lackner, J. R., & Levine, M. S. (1978). Visual direction depends on the operation of spatial constancy mechanisms: The oculobrachial illusion. *Neuroscience Letters, 7,* 207–212.

Sherrington, C. S. (1907). On the proprioceptive system, especially in its reflex aspect. *Brain, 29,* 467–485.

PROSTHESES: AUDITORY

See Cochlear Implant: Technology

PROSTHESES: VISUAL

The goal of *visual prostheses* is to restore visual function in people who are blind by converting visual information into patterns of electrical stimulation that can be used to directly stimulate neurons of the retina or brain. Auditory prostheses have proved remarkably successful in restoring hearing; however, the development of a useful visual prosthesis still faces significant hurdles.

Over the last decade there have been significant advances in a variety of areas that are key to visual prosthetic design, including surgical techniques, electronic miniaturization, electrode design, safe encapsulation of electronics in the body, and our understanding of the neural effects of electrical stimulation. As a result, the field of visual prosthetics has made rapid advances: Currently some dozen academic and commercial research groups around the world are developing devices, and several have implanted test devices into human volunteers.

Almost all current devices in development use a configuration where visual information from a camera is converted into an appropriate stimulating pulse pattern. This pulse pattern is sent to a set of electrodes that then produce percepts (known as *phosphenes*) by electrically stimulating neurons. Prosthetic devices can be categorized on the basis of where the stimulating electrodes are implanted. Possible sites include the retina, optic nerve, visual thalamus, and visual cortex. In this entry, the types of prostheses will be described, as well as other issues related to prostheses, such as eye movements, resolution, field of view, electrode size, and pixilation.

Types

Retina

In recent years, most retinal prosthesis projects have targeted the retina. One reason for this is that surgery on the eye (especially given that patients already are severely blind) is less invasive than brain surgery. A second factor is that at each stage of visual processing, information is represented in a more complex form. Implants that stimulate earlier in the visual pathway need to recreate fewer stages of processing than those implanted in the cortex.

Within the retina, it is possible to implant *subretinally* (placing electrodes in the space between the missing photoreceptors and the choroid) or *epiretinally* (placing electrodes on the vitreous surface of the retina, adjacent to retinal ganglion cell bodies and axons). For both types of implants, the goal is to bypass the missing photoreceptors and create percepts by directly stimulating the remaining cells of the retina (ganglion and/or bipolar cells).

Subretinal implant projects include the Harvard retinal implant project, the Stanford optoelectronic retinal prosthesis, and the retina implant AG project. This last group recently carried out a human trial in which several patients were implanted with test devices for a four-week period. Maintaining a stable array position on the fragile retina during eye movements that can result in roughly 3*g* of accelerative force is likely to be easier for subretinal than for epiretinal implantations, because the array is sequestered between the neural retina and the choroid. However, subretinal implantation is extremely challenging surgically. Other issues include the need to develop a "porous" array so the ability of the choroid to supply nutrients to the neural retina is not restricted, and also the need to provide adequate power to the array without thermal damage.

To date, the only long-term clinical trial of a retinal implant has been an epiretinal device. The first-generation implant, developed by the University of Southern California and the company Second Sight Medical Products, Inc., had 16 electrodes and was implanted in 6 subjects between 2002 and 2004. More recently, this group has begun trials of a 60-electrode implant. Epiretinal implantation is less surgically challenging than subretinal implantation, and there is more potential for heat diffusion into the vitreous cavity. Epiretinal implants generally have the goal of selectively stimulating ganglion cells. Information is carried by a relatively small number of ganglion cells (~ 1 million), as compared to photoreceptors (over a hundred million), or early cortex (hundreds of millions). However, maintaining stability on the retina may be a greater challenge for epiretinal than for subretinal implantation.

One topic of research for both subretinal and epiretinal implants is how to selectively stimulate single cell types. Some selectivity can be achieved by manipulating pulse width: Ganglion cells are more sensitive than bipolar cells when using short pulse durations, whereas bipolar cells are more sensitive than ganglion cells for longer pulse durations. However, there are around 30 different types of ganglion cells, and a similar variety of bipolar cells. Stimulation of each of these cell types is likely to result in a different visual percept, and to date no good method has been developed for selectively stimulating one specific type. An additional concern is that it is not yet clear whether it is possible to consistently avoid stimulation of ganglion cell axons. This last issue is critical because any retinal location near the fovea contains axon fibers projecting to the optic nerve from more distant retinal locations. Stimulation of these axons would be expected to lead to clusters of phosphenes mapping onto unpredictable locations in visual space.

Optic Nerve

Chronic implantation of an optic nerve cuff in a single human volunteer by a group led by Claude Veraart has demonstrated that phosphenes can be elicited at safe charge densities. Advantages of stimulating the optic nerve include having to recreate relatively few stages of visual processing, a potentially less invasive surgical procedure than subcortical or cortical implants, avoiding attachment of the array on the rapidly moving retina, and the restriction of stimulation to ganglion axons. However, the optic nerve is a bundle of fibers, making mapping between stimulation location on the electrode cuff and the position of the elicited phosphene in visual space extremely difficult.

One disadvantage of both optic nerve and retinal prosthetics is that ganglion cells and their axons (which form the optic nerve tract) need to be intact. This restricts use to patients blind due to photoreceptor losses (such as retinitis pigmentosa and macular degeneration) and excludes patients with diseases (such as diabetic retinopathy and glaucoma) that affect ganglion cells or the optic tract.

Visual Thalamus

To date, no visual prosthesis has been implanted in human visual thalamus. However, monkeys trained to make eye movements to the location of light spots will make similar eye movements when stimulated in the thalamus, suggesting that the percepts elicited by electrical stimulation in the thalamus were similar to those produced by discrete spots of light. Advantages of thalamic implantation include targeting the largest possible patient population, and the simplicity of thalamic receptive fields, which are similar to those of ganglion cells. However, in the thalamus there is a separation of magnocellular, parvocellular, and koniocellular pathways. Each of these pathways is thought to mediate different visual functions (e.g., detecting moving stimuli versus fine detail), so being able to stimulate these pathways independently might provide some additional flexibility in representing visual information. The thalamus also contains an over-representation of the central visual field, such that regular spacing of electrodes will create a perceptual spacing highly weighted toward representing the center of gaze.

One challenge is that the thalamus is located deep within the brain. While stimulators for Parkinson's disease are placed in a very similar cortical location, safety requirements are likely to be higher for a visual prosthetic. A second challenge is that thalamic neurons represent the visual field extremely densely. Achieving high sampling densities without tissue damage over years of implantation is likely to prove a significant technical hurdle for implantation within both the thalamus and cortex.

Visual Cortex

There have been several human trials of a cortical prosthesis to date. Experiments using short-term stimulation in the 1970s by Giles Brindley, William H. Dobelle, Daniel A. Pollen, and others demonstrated that electrodes on the surface of the cortex did produce phosphenes. However, required current levels were above safe long-term stimulation limits, electrodes had to be widely spaced to prevent interactions, and patients often saw multiple dispersed phosphenes rather than a single spot of light. In 1996, Edward M. Schmidt implanted a device with penetrating microelectrodes, and found that phosphenes could be elicited at safe current levels. However, the phosphenes produced by a given electrode changed rapidly over time, and the microelectrode array did not maintain a stable position. The main group currently exploring a cortical visual implant is the Utah Artificial Vision group. Although they have no plans to carry out human trials in the near future, the Utah array is currently used in a number of long-term stimulation and recording projects in animal models.

Advantages of a cortical prosthetic include targeting the largest possible patient population, and the accessibility of early stages of processing within the visual cortex (especially those regions of the cortex that represent the center of gaze). Disadvantages include surgical invasiveness, the need to recreate the significant visual processing that occurs prior to the cortex, and the organization of the cortex, whereby a variety of features (color, lines, motion, ocular dominance) are represented within neighboring cells.

Other Issues

As well as the factors previously described, there are certain issues that are relevant to any type of implant, regardless of site.

Eye Movements

In normal vision, the image of the world on the retina shifts whenever we move our eyes. Signals from areas in the brain that control eye movements automatically compensate for this movement on the retina. If the visual information for a prosthetic comes from a head-mounted camera, then changes in eye position will not result in a shift in the retinal image, and the discrepancy between eye movement and retinal information may make objects appear to move whenever the patient moves his or her eyes. One potential solution is to implant a camera within the eye, so that movements of the eye automatically result in an appropriate shifting of the retinal image. Another is to monitor eye movements and shift the visual information provided to the implant appropriately.

Resolution, Field of View, Electrode Size, and Pixilation

There are over 130 million photoreceptors in the healthy retina, whereas prosthetic arrays that are currently in development contain on the order of 1,000 electrodes and cover a field of view of roughly 10° radius.

Pilot data using "simulated prosthetic vision" suggest that useful vision may be achievable with remarkably low resolution and small fields of view. Nonetheless, given a fixed number of electrodes, prosthetic devices need to find the best trade-off between higher resolution and a smaller field of view, as compared to lower resolution and a wider field of view, as shown in Figure 1, parts (a) through (d).

A related issue is that of electrode size. Electrodes must be small enough to support the desired resolution, yet large enough to create phosphenes of the appropriate size. Electrodes that are too small are likely to result in a "pixelated" image that is difficult to interpret. Small electrodes also require high charge densities to elicit phosphenes, limiting the choice of electrode materials. These issues are made still more difficult by the fact that it is not yet clear how electrodes that are close to each other interact electrically and neurally.

It will be difficult to estimate through simulations or animal models what combination of resolution, field of view, and electrode size will be best for useful vision. Optimizing these factors is likely to require a combination of simulation and human trials.

Plasticity

It is not clear to what extent patients will be able to develop the capacity to interpret visual input that does not resemble normal vision. As previously shown, the same quantity of visual information can vary greatly in interpretability. In patients implanted with cochlear prosthetics, there is a significant

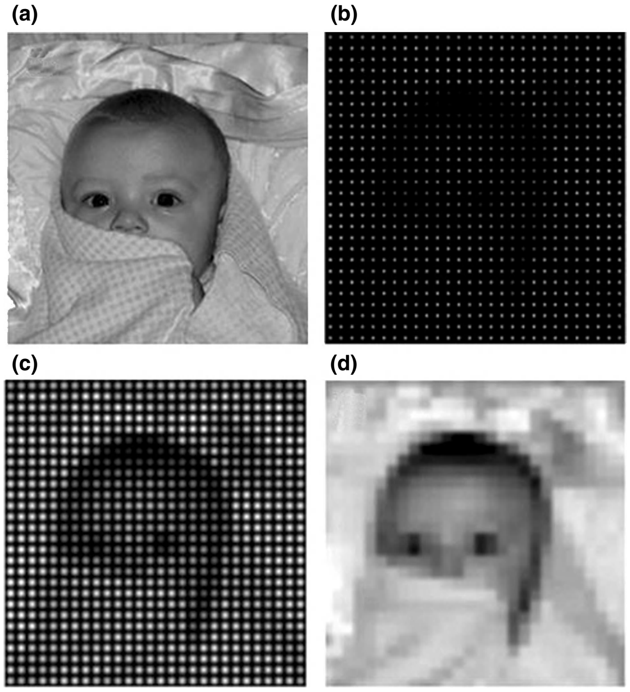


Figure 1 Visual Prosthesis Device and Simulation of Information Created by Phosphenes

Source: Second Sight Medical Products, Inc.

Notes: (a–d) The picture in part (a) is represented by ~1,000 pixels. (b) Simulates a device where the electrodes produce small nonadjacent phosphenes. (c) Simulates larger, nonoverlapping phosphenes, and (d) simulates overlapping phosphenes. These three examples contain identical amounts of visual information but differ greatly in how effectively and naturally that information is represented.

learning period: Speech perception generally improves over a period of several months postoperatively. This learning seems to be most pronounced in those implanted at a young age. Early severe blindness that does not include cortical impairment is extremely rare in the Western world, so most recipients of a visual prosthesis will be implanted as adults. Given that the adult visual cortex is thought to show only limited plasticity, it is likely that patients will make the best use of prosthetic visual information if it can be presented in a way that mimics "normal" vision.

Risk-Benefit Trade-Offs

There is a tendency for any advance in the field of visual prosthetics to be reported with a particular flavor of "hope in the darkness" media coverage. This is a concern because blindness, especially when it occurs in adulthood, sometimes leads to severe depression and a feeling that any sacrifice would be justified for a chance at restored vision. Such attitudes are of concern when selecting patients for clinical trials because they may make it difficult for potential subjects to assess risk-benefit trade-offs. One example of this was the controversial study of Dobelle in 2000, where volunteers paid over $100,000 to be implanted with a prototype device that ultimately failed to restore useful vision.

Looking Forward

This is an exciting time for visual prosthetics: Several groups have safely implanted prosthetics in human volunteers, and many more groups expect to initiate trials in the relatively near future. As the technology and our understanding of the neural effects of electrical stimulation advance, the safety, stability, and effectiveness of visual prosthetic implants is likely to rapidly improve.

Ione Fine

See also Cochlear Implants: Technology; Eye Movements: Physiological

Further Readings

Humayun, M. S., Weiland, J. D., Fujii, G. Y., Greenberg, R., Williamson, R., Little, J., et al. (2003). Visual perception in a blind subject with a chronic microelectronic retinal prosthesis. *Vision Research, 43*(24), 2573–2581.

Pezaris, J. S., & Reid, R. C. (2007). Demonstration of artificial visual percepts generated through thalamic microstimulation. *Proceedings of the National Academy of Sciences USA, 104*(18), 7670–7675.

Schmidt, E. M., Bak, M. J., Hambrecht, F. T., Kufta, C. V., O'Rourke, D. K., & Vallabhanath, P. (1996). Feasibility of a visual prosthesis for the blind based on intracortical microstimulation of the visual cortex. *Brain, 119*(Pt. 2), 507–522.

Veraart, C., Raftopoulos, C., Mortimer, J. T., Delbeke, J., Pins, D., Michaux, G., et al. (1998). Visual sensations produced by optic nerve stimulation using an implanted self-sizing spiral cuff electrode. *Brain Research, 813*(1), 181–186.

Yanai, D., Weiland, J. D., Mahadevappa, M., Greenberg, R. J., Fine, I., & Humayun, M. S. (2007). Visual performance using a retinal prosthesis in three subjects with retinitis pigmentosa. *American Journal of Ophthalmology, 143*(5), 820–827.

Zrenner, E. (2002). Will retinal implants restore vision? *Science, 295*(5557), 1022–1025.

PSYCHOPHYSICAL APPROACH

Psychophysics is the study of sensory, perceptual, and cognitive systems, based on the evidence of human observers making judgments about what they see, hear, or feel. It is the oldest route to understanding the senses, with its roots in classical times and medieval Arab learning, flowering in the hands of 19th century geniuses such as Thomas Young, Hermann von Helmholtz, and Ernst Mach, and with increasing technical sophistication it is a key approach to understanding perception today. Psychophysical methods are also important in the accurate diagnosis of sensory and cerebral disorders and deficits (e.g., color deficiency and visual field loss) and in applications to human engineering problems. This entry draws its examples of the psychophysical approach from the study of visual perception. However, many advances in psychophysical methods, and some of the most sophisticated quantitative applications, have come from work on hearing, and the approaches described here have been effectively applied in the other senses and in studies of multisensory perception.

The term *psychophysics* was coined by Gustav Fechner in 1860 to denote the study of how subjective ("psycho-") attributes of a stimulus relate to its physical properties. For Fechner and a number of his successors, the key question was the way that subjective magnitude of a stimulus (for example, a numerical assessment of a light's perceived brightness) related to its physical intensity (e.g., photometric luminance). However, this question has been a diminishing part of psychophysical research in recent decades.

Psychophysics in a much broader sense has been the use of subjective responses and judgments to make inferences about the structure and function of the visual system and other sensory systems. This approach, for the pioneers of the 19th century, provided their only tool for probing how sensory systems operate. More recently, electrophysiological and neuroimaging methods have allowed researchers to make physical measurements of the brain mechanisms of perception and cognition at work. Psychophysical methods have not been made obsolete by these advances in neuroscience. They remain an approach of great power and precision (and relatively low cost) in analyzing how sensory systems work. More fundamentally, measures of neural activity alone can tell us little about the functional significance of that activity. In contrast, psychophysical judgments of distance, color, or motion must be tapping the systems that we use to gain information about these qualities of the world around us, and so have clear functional validity. Furthermore, the parallel evolution of psychophysical and neuroscience methods has opened the possibility of a new "psycho-physical" relation, directly testing the relationship between physical patterns of activity in the nervous system revealed by imaging and recording methods, and the associated patterns of subjective visual experience. Such linked studies offer a possible way to advance our understanding of the neural basis of consciousness from direct experimental evidence.

Inferring Neural Organization From Psychophysics

Can we use subjective experience to make secure inferences about neural mechanisms? Our visual experience is rich and multidimensional. If we simply ask observers to report what they see, it is often hard to use their responses to test specific hypotheses about how neural signals carry different kinds of visual information. To achieve this, psychophysical experiments usually restrict the observer's report to a clear cut judgment, such as "Did the pattern of dots appear to move up or down?"

Giles Brindley, in an influential discussion, argued that the most secure way to use psychophysical experiments was in "type A" observations; if the observer reported that two stimuli were indistinguishable, it could be inferred, through a plausible "psycho-physical linking hypothesis" that they had an identical effect on the brain. The best example is in one of the oldest types of psychophysical test: color matching. If a mixture of red and green lights appears the same as a yellow light, it is deduced that they are producing identical patterns of response across the light-sensitive cells in the eye. This argument was used by Thomas Young in 1805 to infer that there were just three types of such cells in the eye, and since has been used to characterize precisely how these cone photoreceptors respond to different wavelengths.

A critique of this argument from "type A" observations is that it does not prove at what level in the sensory system the effects are identical. In the color matching case, there are good (physiological) reasons to believe that the effects are matched at the earliest possible level, the cone photoreceptors. However, other examples are more problematic. For example, if a flash of light is delivered to one eye only, observers cannot tell which eye received it. The left and right eyes certainly send distinct signals to the brain, and the difference between them is critical for stereo vision. However, this difference apparently is not accessible at the level of those neural signals that the observer can access when attempting to make the judgment. On one hand, this makes the linking hypothesis more complex and uncertain. On the other hand, it offers the possibility of distinguishing levels and types of sensory processes that are accessible to consciousness from those which are not.

In any case, true "type A" observations are quite rare, and many experiments depend on Brindley's "type B" observations of stimuli, which are judged to be alike in some respect, even though they can be distinguished in others. A common form is the measurement of visual thresholds; a set of stimuli are deduced to have similar effects in the brain because they are all just at the limit of what the observer can detect. So, for example, grating stimuli of different bar widths (spatial frequencies) can be adjusted to find the contrast at which they can just be detected. The gratings do not look identical at this point—fine stripes can be distinguished from broad stripes. Nonetheless, such measurements can be compared with calculations of how effective different gratings would be in activating the receptive fields of visual neurons, and thus provide evidence of the size, shape, and configuration of receptive fields in the human visual pathway.

Analyzing Selective Channels by Psychophysical Interactions

A powerful application of the threshold method is to examine how different stimuli interact. Consider three different grating patterns, for instance, (A) oriented at vertical, (B) oriented at 15° off vertical, and (C) oriented at 60° off vertical. If an observer who has been adapted by exposure to an intense 15° grating (B) shows an increase in the contrast needed to detect the vertical grating (A), this provides evidence that A and B are affecting the same channel. If, however, adapting the 60° grating (C) has no effect on A, this indicates that A and C affect different channels. These channels can be plausibly identified with the orientation-selective neurons that have been measured in the primary visual cortex of cats and monkeys.

Adaptation is only one of several interactions that can be used to analyze neural systems. Another is subthreshold summation (A and B together are easier to detect than either of them separately), and a third is masking (A becomes harder to detect against a strong background of B). All these methods make it possible to psychophysically reveal whether and how a broad perceptual system can be fractionated into selective components, each responding to a limited range of the stimuli (orientations, motions, colors) processed by that system. Similar methods have now also been applied in functional magnetic resonance imaging (fMRI). As with psychophysical responses, the measured signals within a given brain area come from large and diverse numbers of neurons, but researchers can identify the contribution of a particular subset by testing how the signals change when this subset is made less sensitive by selective adaptation.

Psychophysical Analysis of High-Level Processing

Psychophysical methods were first applied to understand how relatively simple, low-level properties of the sensory input—for example, wavelength, contrast, stereo disparity—are encoded. However, with the opportunities given by modern computer graphics to control and manipulate complex displays, similar experimental logic can be applied to complex, global, and high-level aspects of perception and cognition. One example is measurement of how well observers can integrate motion or pattern information over a wide area to a "global" structure. For example, given an array of moving dots with 10% sharing a coherent leftward motion and

90% moving in completely random directions, many observers can correctly report the leftward direction. The percentage of coherently moving dots can be varied to determine a global threshold, which reveals the characteristics of a higher level of visual processing than the ability to identify the local movement of a single dot. Other examples are in face perception, where variables such as the spacing of facial features can be manipulated in computer image processing to test what level of difference can be identified and so determine what properties are most important to human observers in correctly recognizing individual faces. These approaches are proving important in studying clinical groups. One example is children with a range of disorders (e.g., autism and dyslexia) who are impaired more in global motion integration than in a comparable task that depends on integrating static pattern information. Another example is the evaluation of individuals for whom brain damage leads to difficulty with faces; they may recognize individual facial features but be insensitive to the configuration that they form in a person's face

Statistical Decisions in Psychophysical Judgments

Like many measurements in science, psychophysical judgments are not exactly the same on every repetition, but show random variation or "noise." Some of this noise is external to the observer—for example, statistical fluctuations due to the quantum nature of light place a limit on our ability to detect small variations in light intensity. Other noise is internal—due to variability in neural events or fluctuations in the observer's attention. Whatever the source of noise, its constant presence means there is no such thing as a fixed threshold at which a stimulus goes from being invisible to visible. Rather, there is always a finite *psychometric function*. This function is a curve plotting how the percentage of occasions that an observer reports that he or she saw a flash, for example, increases smoothly with the amount of light energy in the flash. Some conventional value (e.g. 75% correct detection) is taken to define the "threshold" amount of energy.

This fundamentally statistical nature of psychophysical judgments has several important consequences. First, it means that in making such a judgment, we are making a statistical decision under uncertainty. Variable sensory events occur, and we have to decide whether they reflect the actual occurrence of a stimulus. This is best done by setting an internal criterion so that whenever the event is stronger than this criterion, we judge that the stimulus was present. On this view, the "threshold" does not simply reflect the sensitivity of our sensory system to faint stimuli, but also the way that we operate our decision processes. *Signal detection theory* provides quantitative methods to design and analyze psychophysical experiments that can separate sensory sensitivity from decision criteria—the latter are affected by such things as the observer's caution, her degree of confidence, and the relative value to her of correctly spotting every stimulus versus avoiding any "false alarms." It emphasizes the particular value of "forced-choice" experiments: If an observer, instead of being asked "Did you see a flash? Yes or no?" is told "There will be two time intervals—tell me whether you think the flash occurred in the first interval or the second," the results will not be affected by any changes in her criterion for reporting a flash.

Second, the statistical approach allows perceptual performance to be evaluated against theoretical principles of how information in the stimulus can optimally be used. Any noise external to the observer (i.e., due to variations in the stimulus) sets a fundamental statistical limit on how weak a stimulus can be detected. In some cases (e.g., the quantum fluctuations that limit the detection of a very weak flash of light) human vision approaches this theoretical limit. In other cases, the efficiency of human visual processes relative to this limit can be calculated. This provides a measure of performance that can be compared across different visual tasks to examine which tasks the human brain is best adapted for.

The statistical approach also makes it possible to test whether perceptual information is efficiently combined, both between senses and with information coming from prior knowledge. Bayes' theorem provides a measure of the optimal combination of different sources of information. The Bayesian approach to perceptual judgments has provided added quantitative rigor to test the old idea that perception is partly determined by expectations based on our knowledge of the world.

Third, the statistical nature of detection means that the idea of "subliminal perception" has to be treated with great care. Two different tasks with the same stimuli may encourage the observer to apply different criteria for detection; if, as a result she shows effects in task B of a stimulus she would not report in task A, this does not necessarily imply any mysterious ability to "respond to unseen stimuli."

Fourth, it means that psychophysical responses to randomly varying patterns can be a powerful tool for analyzing perceptual systems. This is the basis of the relatively new method of "classification images." As an example, suppose we take two slightly misaligned vertical dark bars and ask which part(s) of the display observers are using to judge the direction of misalignment—a "vernier" task. If we blanket the display in different random computer-generated texture noise on each presentation, we can find the average noise pattern on trials where the observer judged "misaligned left" compared to those judged "misaligned right." This averaged noise pattern will be dark in those regions that are most important in affecting the judgment, and so define the "filter" or "receptive field" that the observer is using within her visual system to perform the task.

Cues and Perceptual Learning

The skill of experimental design in psychophysics using the forced-choice method lies in devising stimuli whose only systematic difference—the "cue"—must be detected by the mechanism that the experimenter wishes to study. If successful, there is no opportunity for the observers to "cheat" by using other cues or mechanisms, either consciously or unconsciously. (For example, if the process under study is the ability to discriminate different speeds of motion, the stimuli should not allow the observer to systematically distinguish the moving stimuli by how long they last or what distance they travel.) For the observers' part, they have to recognize that cue and use it as the basis for their decisions. Any stimulus will activate many millions of different neurons in the visual parts of their brains, and as previously discussed, this activity will include random variations or noise. Somehow, the observers have to pick out which part of this noisy torrent of neural information is relevant to their task. It is not surprising, then, that psychophysical experiments generally require some pretraining, nor that observers' performance almost always improves markedly with practice, over hours and even days. This improvement is an example of perceptual learning—a phenomenon that occurs in everyday life as we learn to distinguish the characteristic forms and brushstrokes of paintings by Monet from those by Renoir, or the sound of our vacuum cleaner when it is working properly from its sound when the drive belt is broken. It raises the question of how far psychophysical experiments used to test a hypothesis in a specific experiment are testing the sensory system in its "natural" condition. The answer must be that we have certain basic sensory machinery, but we are continually marshaling it in different ways to meet the demands of specific tasks. When inferring the mechanisms that underlie psychophysical performance, we have to include both basic processes that are used for many different tasks, both inside and outside the laboratory, and specific strategies that the observer has developed to use those processes efficiently in the context of the experiment. For example, when a masking method is used to analyze auditory channels, observers can sometimes achieve the most sensitive threshold not by using the channel that would normally detect a tone of a particular frequency, but by "off-frequency listening" using a nearby channel that is less affected by the mask. The results can only be understood by taking into account both the characteristics of the channels and the cues that may be provided using various different channels as to whether the tone is present.

Animal and Infant Psychophysics

Because psychophysical experiments typically use verbal instructions ("press key 1 if the flash appears in the first interval and key 2 if it appears in the second interval"), they might be thought of as an approach to sensory mechanisms specifically for human participants. However, given that learning is an essential part of the participant's preparation, psychophysical experiments can also be designed using our knowledge of operant learning methods for many nonhuman species. Color matching in goldfish, dark adaptation in pigeons, and depth perception in kittens are the subjects of three classic experimental programs based on training methods that harness characteristic behaviors of these species to controlled stimulus variations.

Such experiments have shown that the close anatomical homologies between the human visual system and that of macaque monkeys (photoreceptor types and packing density, layout of the visual pathway) are reflected in similar psychophysical performance. The color matches, visual acuity, contrast thresholds for different gratings, and stereo vision of macaques trained in these psychophysical tasks are essentially identical to those of human observers. These functional homologies mean that a range of physiological evidence from the macaque visual system can be used in understanding human visual perception (although we must also be aware of ways in which the brains of the two species may have diverged in evolution).

Another type of participant that cannot follow verbal instructions is the young human infant. Nonetheless, many of the stimulus manipulations used in psychophysics can also be used in preferential looking studies, where infants' detection is inferred from their pattern of eye movements toward stimuli that they can detect—effectively psychophysics with a nonverbal response.

Parallel Psychophysical and Neural Measurements

Behavioral methods with macaques have also made it possible to make psychophysical measurements and neurophysiological measurements at the same time on the same individual. In now classic experiments, William Newsome, Ken Britten, and colleagues measured how single neurons in the middle temporal (MT) area of monkey cortex responded to global motion patterns at different coherence levels, while simultaneously the monkeys made judgments, which they signaled by their eye movements, of the direction of movement. They found that some individual cells show similar thresholds to the animal's behavioral judgments for the same presentations, implying that individual, or very few, cells could provide the information on which the monkey based its judgments. This was a striking result because it might be expected statistically that the judgment would benefit from combining independent information from many thousands of cells. It raises intriguing questions for future research, first of a technical nature about the statistical independence or otherwise of different neurons, and second, of a more fundamental nature about how the monkey (or the human) identifies which neurons are reliably informative and which are not.

Issues about which neurons determine the contents of visual perception have also been highlighted by using behavioral judgments and neuronal measurements in the study of binocular rivalry. When monkeys are presented with incompatible images to the two eyes (e.g., red vertical stripes to one eye and green horizontal stripes to the other), their responses show, like human observers, that they perceive each of the two images as dominating alternately. Work by Nikos Logothetis and colleagues has shown that some neurons in early visual areas (V1–V4) turn on and off their responses in line with what the monkey is "reporting"—but other neurons behave independently of, or even opposite to, the perceptual effect—for example, responding to the green horizontal stripes at a time the monkey is seeing red verticals. At higher levels in the temporal lobe, neural responses are more consistently uniform with what the monkey is apparently seeing. This work provides some of the most detailed evidence to date that perceptual experience is associated with a subset, but only a subset, of neural activity in visually related brain areas.

As functional neuroimaging methods have become more refined, it has become possible to perform analogous experiments in humans, exploring the relationship between psychophysical judgments and brain activity, albeit at a less specific level than single neurons. For example, activity in the "face area" in the fusiform gyrus has been shown to be modulated according to whether the face is currently visible or suppressed in binocular rivalry. Another line of neuroimaging research has shown different specific patterns of activation in the visual cortex according to whether a given visual stimulus was detected, undetected, or falsely reported on a particular presentation. Progress with these kinds of experiments suggests that sensory science is entering an era when psychophysical methods and measurements of neural activity are not just complementary, but are fully integrated approaches to understanding how perceptual systems work and how they underpin human conscious experience.

Oliver Braddick

See also Bistable Perception; Color Mixing; Consciousness; Decision Making, Perceptual; Face Perception; Infant Perception: Methods of Testing; Motion Perception; Perceptual Learning; Physiological

Approach; Psychophysics: Detection; Selective Adaptation; Signal Detection Theory and Procedures; Visual Processing: Extrastriate Cortex; Visual Processing: Primary Visual Cortex

Further Readings

Barlow, H. B. (1995). The neuron doctrine in perception. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 415–435). Cambridge: MIT Press.

Brindley, G. S. (1960). *Physiology of the retina and visual pathway.* London: Edward Arnold.

Carpenter, R., & Robson, J. G. (Eds.). (1998). *Vision research: A practical guide to laboratory methods.* Oxford, UK: Oxford University Press.

Eckstein, M. P., & Ahumada, A. J., Jr. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision, 2*(1), i. Retrieved from http://www.journalofvision.org/2/1/i/. DOI:10.1167/2.1.i

Goldstein, E. B. (2005). Cross-talk between psychophysics and physiology in the study of perception. In E. B. Goldstein (Ed.), *Blackwell handbook of sensation and perception* (pp. 1–23). Oxford, UK: Blackwell.

Levine, M. W. (2000). Psychophysics. In *Fundamentals of sensation & perception* (3rd ed., chap. 2, pp. 5–27). Oxford, UK: Oxford University Press.

Logothetis, N. K. (1998). Single units and conscious vision. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 353,* 1801–1818.

Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience, 21,* 227–277.

Teller, D. Y. (1980). Locus questions in visual science. In C. S. Harris (Ed.), *Visual coding and adaptability* (pp. 151–176). Hillsdale, NJ: Lawrence Erlbaum.

Watt, R. J. (1991). Psychophysics and image algebra. In *Understanding Vision* (chap. 9). London: Academic Press.

Watt, R. J. (1991). Studying human vision. In *Understanding Vision* (chap. 8). London: Academic Press.

PSYCHOPHYSICS: DETECTION

A detection threshold is the minimum stimulus strength that can be perceived. Even with such a humble definition, the measurement of detection thresholds has become a golden, powerful tool for psychophysicists, but it is elusive to pinpoint. Some examples of visual detection tasks discussed in this entry are:

- *Detection of a tumor.* In order to diagnose cancerous tumors, radiologists must learn to detect their presence in a complex X-ray.
- *Detection of a feature.* The ability to detect a gap in the letter C can indicate whether new eyeglasses are necessary, or whether, even with new eyeglasses, one is fit to drive a car.
- *Detection of a needle in a haystack.* Many real-life detection tasks are similar to finding Waldo in a crowd or a needle in a haystack. Careful analysis of how observers detect small targets in a cluttered environment can provide insight into whether brains can process many stimuli simultaneously and into how uncertainty limits visibility.
- *Detection of a thin line.* The timing of certain Jewish and Islamic religious holidays is based on the earliest detection of a new crescent moon by religious authorities using unaided vision.
- *Detection of a star.* A ritual requiring the bride and groom to detect the double star pair Mizar and Alcor in the handle of the Big Dipper is a part of many Hindu wedding ceremonies. Successful sighting portends a successful marriage.
- *Detecting a minimal shift in the position of a target feature.* This can be understood in terms of detecting a thin line, like the fourth item in this list, and as discussed in the following paragraphs.
- *Detection of a glitch in a movie.* Twenty years ago, video engineers were convinced that high resolution digital TV was impossible because the bandwidth they believed was necessary was hundreds of times larger than was feasible. However, vision scientists showed that limits on detection of tiny features, especially in clutter, enables dramatic image compression, making high definition digital TV possible.

We will come back to some of these examples later. But first we must take a detour into how detection thresholds can be measured.

A common method for measuring detection thresholds is the “Yes–No” method. In this method, the observer is shown, in random order, repetitions of a variety of stimulus intensities, ranging in strength from zero to clearly detectable. The task is to say yes or no according to whether a target was seen. The outcome of the experiment can be presented as an S-shaped plot (Figure 1a) showing the percent of times the person said yes, as a function of signal strength (*s*). The vertical axis is often called the hit rate (p_{hit}). This plot of the psychological response to a physical stimulus, called the psychometric function, is fundamental to the field of study called psychophysics.

The lowest point on the curve, called the false alarm rate ($p_{false\ alarm}$) represents the probability of the observer saying "yes" even when no stimulus is present. The value of $p_{false\ alarm}$ (= 6.7% in the Figure 1a example) can be easily shifted by the observer's intentions. We now consider three possible definitions of detection threshold, shown as dots on the plot.

The *null hypothesis threshold* is the value of *s* where p_{hit} first significantly deviates from its level at $s = 0$. The Figure 1(a) example has $p_{false\ alarm} = 6.7 \pm 2\%$, where 2% is our assumed standard error of $p_{false\ alarm}$. For hypothesis testing, it is common to use the z = 2 standard error point that is at 6.7% + 2 × 2% = 10.7%, as shown on the plot. By consulting a table of z-scores, one learns that for z = 2, there is only a 2.3% chance that a hit rate of 10.7% could have occurred by chance, even if there had been no signal. For those unfamiliar with z scores, it is similar to the IQ scale given by z = (IQ – 100)/15. Thus, an IQ = 130 gives z = 2, which means 2.3% of the population has an IQ > 130. This definition of detection threshold has the problem that if one had increased the number of presentation trials used for measuring the detection threshold, the standard error would have decreased below 2% (more trials means greater accuracy), and the threshold would have decreased. Psychophysicists prefer a threshold definition that is independent of the number of trials.

The *midpoint threshold* is the value of *s* at the psychometric function's midpoint, which is the average of the upper and lower asymptotes (p_{hit} = 53.3% in Figure 1a). The problem with defining threshold at either the midpoint or at a fixed p_{hit} = 50% is that it doesn't adequately take into account the observer's bias. Consider the case of radiologists examining X-rays for signs of cancer. In cases where biopsies are relatively noninvasive, a moderate false alarm rate, as in Figure 1(a), is reasonable ($p_{false\ alarm}$ = 6.7% means that about 1 in 15 people without a tumor would get a biopsy). However, for highly invasive biopsies, a lower false alarm rate of say 0.2% (1 in 500 healthy people being biopsied) may be appropriate. With this much stricter criterion, the entire

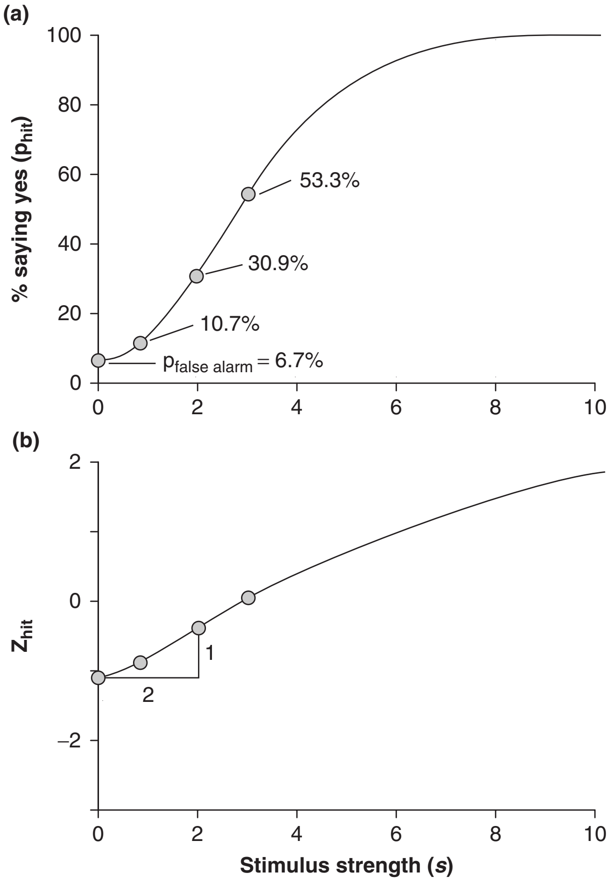


Figure 1 Psychometric Functions

Notes: Figures 1(a) and (b) show the psychometric function for saying "yes, I see the target" as a function of stimulus strength on the abscissa. Part (a) shows the probability (%) of saying yes (the hit rate, p_{hit}) and part (b) is the z-score transform of the upper panel. The leftmost dot on each curve shows the false alarm rate, which is the hit rate for a nearly zero stimulus strength. The other three dots correspond to three definitions of threshold, as discussed in the text. The one that is relatively independent of subjective bias corresponds to the middle of the three dots in the lower panel: The increase in stimulus strength needed to increase the z-score effect by one. Thus, in this example, the detection threshold would be 2 units of stimulus strength.

psychometric function will shift rightward, and thus the midpoint definition would correspond to a larger stimulus strength, making it appear that the actual threshold is higher. However, psychophysicists prefer to characterize intrinsic ability to detect a target, independent of the observer's cost-benefit bias. Thus, we can see that, because of its variability as a function of observer bias, the midpoint definition cannot identify an intrinsic threshold.

The *signal detection theory threshold* is the value of s at which the p_{hit} z-score (z_{hit}) increases by one. About 60 years ago, engineers charged with detecting signals in radar images developed a mathematical/statistical approach to the detection threshold problem called signal detection theory (SDT), now commonly used by psychophysicists. The main step in this approach involved shifting from probability (Figure 1a) to z-score (previously introduced) for plotting the psychometric function, as shown in Figure 1(b). SDT defines threshold as the shift in s that is needed to increase z_{hit} (the vertical axis in Figure 1b) by one above its value at $s = 0$. The triangle of Figure 1(b) shows that when the stimulus strength increases from $s = 0$ to 2, z_{hit} increases by one (going from −1.5 to −0.5, corresponding to probabilities of 6.7% and 30.9% as shown in Figure 1a). Therefore, for our example, the detection threshold is $\Delta s = 2.0$. This SDT definition avoids being affected by an observer's criterion because according to SDT (and validated to a good approximation by many experiments), a criterion shift leaves the shape of the psychometric function unchanged (with z_{hit} as the ordinate) and merely shifts it vertically. In some domains of psychology, the amount of increase in z_{hit} over its value at $s = 0$ is called d' (d prime). In other domains of psychology, it is called "effect size." For those with a statistics background, it may be useful to point out that the difference between using the null hypothesis definition (a) and the SDT definition (c) is the difference between using the standard error (t-test) versus the standard deviation (effect size) as the metric for defining detection threshold.

It is useful to mention the main alternative to the Yes-No method called the *two-alternative forced choice* (2AFC) *method*. On each 2AFC trial, two stimuli are shown, the target and the reference, in random order, and the observer is asked which interval had the target. This method has the advantage of being less dependent on subjective criteria. However, Jacob Nachmias showed that it is often the case that thresholds in discrimination tasks are 30% higher when targets are in the first interval, thereby calling into question assumptions of the 2AFC method.

We now examine some of the mathematical and psychophysical underpinnings for the visibility of three of the simpler items mentioned at the beginning: a star, a new crescent moon, and a position shift.

The threshold for seeing a star (like the marriage double star pair, Mizar and Alcor) under conditions of dark adaptation was carefully studied by Selig Hecht, Simon Shlaer, and Maurice Pirenne using the Yes-No method. Based on the shape of their psychometric functions, they concluded that the detection threshold of a point of light was six photons absorbed by rods. Their calculation made use of properties of photon noise whereby the steeper the psychometric function (Figure 1a), the more photons were needed for detection. The significance of Hecht's finding is that a psychophysical procedure was used to determine a basic fact about the activation of photoreceptors, which then had to be explained by physiologists (and it took a while!). This application of detection emphasizes that measures of detection can help us uncover basic mechanisms.

We next show how knowledge of the detection threshold of a thin line gives insight into the visibility of a new moon. We start with the finding by Selig Hecht and Esther Mintz that the detection threshold for seeing a long, thin line is about 2% min (percent of contrast per minutes of arc). This threshold means that as long as the line width in minutes of arc times the line contrast in percent exceeds 2% min, the target is detectable. For a short line 15 min long, the detection threshold is about twice as large or around 4% min. We can now relate this human threshold to the topic of the visibility of the crescent moon. The crescent becomes just visible about one day after the new moon, when its width at the widest point is about 1 min and its effective length is about 15 min. Based on the previous discussion of human line thresholds, the new moon's contrast at sunset should be about 4%. This analysis provides a beautiful example of how an understanding of detection thresholds can be relevant to deciding when the Islamic year begins.

We can also predict position shift thresholds using earlier information. Let's start with a sharp edge made by two abutting sheets of white paper on a dark background (a 100% contrast edge). Shifting the left sheet slightly upward produces an offset that is the same as adding a thin line of 100% contrast to a straight edge. The Hecht and Mintz line detection threshold of 2% min implies that a thin line of 100% contrast (smoothly matching the edge) and 0.02 min width would be just visible. This angular offset can be obtained by viewing a 4 millimeter (mm) jag in the paper from a distance of 1 kilometer (km), and this is the angular offset that is at the detection threshold of a human being (Dennis M. Levi) as reported in Guinness World Records.

Although the previous discussions concerned tasks for detecting visual targets on either uniform or structured backgrounds, the same approaches and same mathematics applies to all sensory modalities, not just vision. The generality of the previous approaches and ways of thinking, whereby almost any discrimination task in any sensory modality can be reduced to a simple detection task, is one of the beautiful aspects of the field of psychophysics.

Stanley A. Klein and
Olga Eizner Favreau

See also Psychophysical Approach; Signal Detection Theory and Procedures; Visual Acuity

Further Readings

Hecht, S., & Mintz, E. U. (1939). The visibility of single lines at various illuminations and the retinal basis of visual resolution. *Journal of General Physiology, 22,* 593–612.

Hecht, S., Shlaer, S., & Pirenne, M. H. (1942). Energy, quanta and vision. *Journal of General Physiology, 25,* 819–840.

Klein, S. A., & Levi, D. M. (1985). Hyperacuity thresholds of one second: Theoretical predictions and empirical validation. *Journal of the Optical Society of America. A, Optics and Image Science, 2,* 1170–1190.

Nachmias, J. (2006). The role of virtual standards in visual discriminations. *Vision Research, 46,* 2456–2464.

Q

QUALIA

Perception and thought are often, although not exclusively, concerned with information about the world. In the case of perceiving, unlike thinking, it is widely believed that there is an additional element involved, a subjective feeling or, as it is often put, something that it is like to be perceiving. *Qualia* are these characteristic feelings that accompany perceiving. One motivation for the idea that we experience qualia is that there is a clear difference between seeing a red tomato and thinking that a tomato is red and that the difference has to do with some extra element present in the case of seeing that is absent in the case of thinking. Philosophical attempts to understand qualia and their place in the world have played a central role in recent debates about the nature of mind and its place in the world. Before getting to those debates, this entry takes a more detailed look at the distinction between the content of perceptual experiences, what they tell us about the world, their qualitative or phenomenal character, and what it is like to experience them.

Representational Content and Phenomenal Character

Perception informs us about the nature of things in the world around us. I smell the lilacs in my backyard and see the orange flowers of the marigolds. I feel the wind on my face and hear the blaring horn of a fire truck passing through a nearby intersection. In all of these cases, the things I am smelling, seeing, feeling, or hearing are ordinary objects such as bushes, flowers, air, and large motor vehicles. These things may or may not have the properties I perceive them to have, but it is the flowers that may or may not be orange and a large motor vehicle that may or may not be emitting loud noises. In the common jargon of philosophy, we can say that my perceptual experience *represents* the marigolds to be orange and the fire truck to be loud. Correspondingly, we can characterize the (representational) *content* of my experience as including the orangeness of the flowers and the loudness of the truck. One aspect of the content of my experience of looking at a marigold is similar to an aspect of the content of my experience of an orange. Both experiences have as a part of their content that something is orange.

When I look at a marigold, I will, in many circumstances, visually represent it to be orange. In addition, there is something it is like to see (or appear to see) a marigold as orange. *Seeing* a marigold to be orange is different from *reading* that marigolds are orange, and at least part of that difference derives from a difference in what it's like to visually experience something as orange compared with the experience of reading the word *orange*. Again using the jargon of philosophy, the experience of seeing a marigold as orange has a particular *phenomenal character*. Just as the content of the experience of seeing a marigold is similar in some ways to the content of seeing an orange, the phenomenal character of the two experiences is also similar. Experiences that are similar in what it is like to have them, in their phenomenal character, are similar in their experienced qualia, whereas experiences that differ in their phenomenal character are different in their experienced qualia.

The phenomenal character of an experience is, at least in concept, distinct from its representational content. We have picked out the two features of an experience in different ways: what it's about versus what it's like to have it. We can all agree that perceptual experiences have both a content and a phenomenal character and set aside for a moment what the relationship is, if any, between these two features of experience.

As we have seen, qualia are just those properties that explain the similarities and differences in what it is like to undergo perceptual experiences. At this point, the philosophical controversies begin. These controversies revolve around two fundamental questions: What are the bearers of qualia and what kind of properties are qualia?

Are Qualia Nonphysical?

Philosophers have extensively debated whether qualia are physical properties or can be reduced to physical properties. Although just what concept of a physical property is being relied on and what it means to reduce a property are often obscure, the essential ideas behind the controversies can be explained without entering into these complexities. The basic method of argument is by consideration of hypothetical cases. The most straightforward kind of case involves the possibility of a being just like oneself in all bodily characteristics, reaction to external stimuli, and other behavior that does not experience qualia. Call such a hypothetical duplicate of me a zombie because it is just like me except that there is nothing that it is like to be it. Our question then is this: Are zombies possible or, in other words, is it possible to duplicate my body and with it all my patterns of behavior without also duplicating the phenomenal character of my perceptual experiences? When looking at an orange marigold, my zombie duplicate will describe it just as I would, discriminate its color from those of adjacent flowers just as I would, and would in general be indistinguishable in behavior and physically (or physiologically) measurable characteristics. Because zombies are exactly like us in all physical and physiological respects but differ from us in lacking qualia, their possible existence would show that qualia are not among our physical characteristics and don't depend on those characteristics either. That the conclusion that qualia are distinct from and independent of all physical properties follows from the assumption that zombies are possible is relatively uncontroversial. The principle debates revolve around the question of whether zombies are genuinely possible. The chief argument for the possibility of zombies is based on the apparent fact that we can coherently imagine zombies and the assumption that anything that is coherently imaginable is possible. Although it would be preferable to base our conclusion that our physical nature either does or doesn't explain all of our mental lives on observable facts, rather than imagined possibilities, it has proven difficult to find facts that would settle these disputes.

Another much discussed argument for a somewhat weaker conclusion is the knowledge argument. The heart of the knowledge argument is another hypothetical scenario. In this scenario, which takes place in the distant future, we are asked to imagine a person, call her Mary, raised from birth in a completely black-and-white environment, a person who has never had the experience of seeing red (or green or any other color except black, white, and grey). In addition, Mary knows everything there is to know about the physical causes of color, the processing of color information in the brain, and everything else that is scientifically relevant to understanding color experience. As it is usually put, Mary knows all the physical facts (and true theories) relevant to the experience of seeing red. Suppose that Mary is for the first time in her life exposed to a red thing (in good light and at an appropriate distance, etc.). When she first has the experience of the color red, would she learn something that she didn't already know? If she does learn something, then it would seem that facts about qualia are not counted among the physical facts (because she already knew all those facts). If facts about qualia are not physical facts (and not derivable from physical facts), then it would seem that the qualia themselves must also be nonphysical. If, as is commonly done, we assume that physical happenings have only physical causes, then it also follows that qualia, being nonphysical, have no effect on the physical world. In the jargon of philosophy, qualia are epiphenomenal, dependent on the physical world, but not the same as any part of it and powerless to produce any physical effect. One implication of the epiphenomenal nature of qualia is that they can't be the causes of behavior. Mary may learn something new when she experiences the color qualia for the first time, but those qualia can't be the cause of her behavior, including her saying that she has learned something new. Speaking involves physical movements, and those movements can't be caused by epiphenomenal qualia.

One striking feature of both of these arguments is the degree to which they depend on hypothetical scenarios and particular judgments concerning those scenarios. The centrality to the anti-physicalist argument of the intuitive judgment that zombies are possible or that Mary learns something new raises important questions about methodology in philosophy. The status of intuitive judgments of this kind and their proper role in philosophical theorizing has become an important debate in its own right.

Content and Qualia Again

One possible alternative to anti-physicalism concerning qualia is the thesis that the phenomenal character of an experience and its representational content are the same thing differently described. According to this view, called intentionalism or representationalism, there are not distinct characteristics of perceptual experiences, quale and content, but only a single characteristic with two different names. Just as there are not two distinct properties, being water and being H_2O, just a single property with two different names, there may not be two distinct properties of the experience of a marigold, what it's like to have it and what it says about the world. If we could have an understanding of how perceptual experiences come to be about the world in physical terms, then we could also understand the qualia possessed by those experiences in physical terms because the two aspects of the experience are one and the same. One attraction of this view is that it takes two problems, one of which is particularly intractable, and reduces them to a single, one hopes, more tractable problem. Another, much discussed attraction of this view is that it provides a compelling explanation of an important feature of perceptual experience. When I look at an orange marigold, I would not normally describe my awareness as being of two distinct things, the quale of orangeness and the orangeness of the marigold. Although we can distinguish in concept between what it's to like to have the experience and what the experience tells about the world, there are not actually two distinguishable elements of the experience. Perception just seems like an awareness of things in the world, and the usual descriptions of experiences are in terms of what they are about, rather than how they feel. We can't easily find in our experience qualia distinct from content. If the phenomenal character of an experience is just its content, then we have an explanation for why it is so difficult to distinguish the two aspects of experience. It's like trying to separate the water from the H_2O in my glass.

Intentionalism, by identifying qualia and content, is committed to it being impossible to vary one without varying the other. If water is identical to H_2O, then any manipulation that changes the boiling point of water must also change the boiling point of H_2O, and a similar logic will apply to representational content and phenomenal character. The most common objection to intentionalism is that it is possible to independently vary the content and qualia associated with an experience and that this possibility is shown by consideration of certain types of inverted spectrum scenarios. Again, despite various complexities, the basic idea behind the inverted spectrum is quite simple. It seems possible that there could be two people, Fred and Mary, who when looking at ripe tomato on the vine have experiences with inverted qualia. That is to say, that the quale associated with Fred's tomato experience is the same as the quale associated with Mary's leaf experience and vice versa. It also seems possible that Fred and Mary don't differ in their ability to name the colors, or discriminate the colors, or to pick the fruit rather than the leaves from the plant. Given this identity in their abilities to use perception to interact with the world, it seems that their respective experiences must have the same content, be about the same things. We here have a case of difference in qualia with sameness in content. If such a case is genuinely possible, then intentionalism has been refuted.

Qualia in Science

Set aside the anti-physicalist arguments, and take the inverted spectrum argument against intentionalism seriously. We are left with the view that qualia and content are distinct features of perceptual experience and that the quale associated with an experience is identical to or reducible to the physical embodiment of that experience. If qualia are to be explained in terms of brain features, then neuroscience is the proper place to look for an account of the phenomenal character of experience. This expectation, however, may be defeated by the same kinds of considerations that raise difficulties for intentionalism. Because there are no obvious connections between the physiological properties of neurons and the phenomenal character of experience, it is possible to construct inverted spectrum scenarios that conflict with any attempt at physiological explanation of qualia. Why should activity in any particular neuron be associated with the quale of green rather than the quale of red? We could discover correlations between neural activity and qualitative character, but such correlations would not themselves answer the why question. (This is one moral that could be drawn from considering the case of Mary the color scientist.) This line of reasoning leads to the assertion of an explanatory gap: No amount of information about how the brain works will allow us to understand why our experiences have the qualia that they do. To assert the existence of an explanatory gap is different from asserting that qualia are not physical. Here, what is being questioned is our ability to scientifically study qualia, not their status as physical or not. From this point of view, qualia may be physical (or not) but their relation to the physical properties discoverable by science is unknowable.

David Hilbert

See also Mary the Color Scientist; Philosophical Approach; Private Nature of Perceptual Experience

Further Readings

Block, N. (1990). Inverted earth. In J. Tomberlin (Ed.), *Philosophical perspectives* (Vol. 4, pp. 53–79). Atascadero, CA: Ridgeview.

Dennett, D. (1988). Quining qualia. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science.* Oxford, UK: Clarendon.

Harman, G. (1990). The intrinsic quality of experience. In J. Tomberlin (Ed.), *Philosophical perspectives* (Vol. 4, pp. 31–52). Atascadero, CA: Ridgeview.

Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly, 32,* 127–136.

Ludlow, P., Nagasawa, Y., & Stoljar, D. (2004). *There's something about Mary: Essays on phenomenal consciousness and Frank Jackson's knowledge argument.* Cambridge: MIT Press.

Palmer, S. E. (1999). Color, consciousness, and the isomorphism constraint. *Behavioral and Brain Sciences, 22*(6), 923–943.

Tye, M. (2000). *Consciousness, color, and content.* Cambridge: MIT Press.

R

Rapid Serial Visual Presentation

Producers of music videos have often been known to include sequences where the images change 10 times per second or more. Despite the fact that each image is only present for about a 10th of a second, we typically have the impression of having seen the image clearly, even if we were asked to provide a complete list of all the images that were shown. This example leads to the topic of this entry, *rapid serial visual presentation* (RSVP), which demonstrates the remarkable ability of the human visual system to process rapidly presented visual information.

The use of RSVP as a research tool started in the early 1970s. In typical experiments, stimuli such as letters, digits, words, or pictures are shown in rapid succession at rates of up to 20 per second. In some cases, a characteristic such as color is used to specify the target, and the subject's task is to identify the target. One of the original motivations was to determine to what extent the natural rate of exploration of a visual scene using saccadic eye movements (about three fixations per second) is limited by the processing speed of the visual system, or whether the time taken to program and execute the eye movements is a significant factor. The fact that text can be presented at rates well above the three fixations per second typical of reading seems to imply that the bottleneck is probably not in the visual systems ability to process information. In recent years, there has been a great deal of interest in the possibility of using RSVP as a way of presenting text information on small displays, such as the ones used for mobile phones and personal digital assistants (PDAs). By automatically presenting a text as short strings of characters at a rate that can be adjusted to the users reading speed, even long documents can be read using a small display. Indeed, there is now a wide range of software applications available that allow users to try reading text without moving their eyes.

In the lab, RSVP techniques have been used to study a number of intriguing phenomena. These include *repetition blindness,* which is the observation that when the same target item appears twice in the same sequence, we often fail to notice the second presentation. Another intensively studied phenomenon is the so-called *attentional blink,* which occurs when two different targets occur close together in time (200–500 milliseconds, ms): Detection of the first target is followed by a transient drop in the ability to detect the second one. One might think that the first stimulus is somehow masking processing of the second one. However, it has recently been found that if the second target appears immediately after the first, performance is virtually intact (a phenomenon known as *lag-1 sparing*). This implies that the first target does not saturate visual memory, and that more than one target can be processed in one packet, as long as there is no distractor between them. One possibility is that the distractor appearing immediately after the first target segments the stream into chunks that are stored separately.

Although the earliest work with RSVP typically used letters and words, by the mid-1970s, Mary Potter demonstrated that even complex photographs of natural scenes could be processed efficiently in RSVP streams. She showed strings of images at different rates and asked subjects to respond if a particular target stimulus was present in the stream. At roughly 10 frames per second, subjects could reliably report if a target stimulus was present in the sequence, even when the target was only specified verbally ("a boat," or "a baby"). Interestingly, the presentation rate needed to be slowed down considerably for the subjects to be able to remember what they saw. This was tested by giving them a recognition memory test, and after each trial in which they were shown an image, they were asked to report whether they had just seen it or not. This suggests that the time courses for processing an image and storing information about the image are quite different. More recent work has confirmed this by showing that even though images can be accurately identified when presented at high rates, there does not appear to be any automatic storage going on. In many situations, seeing a stimulus once will produce *priming*, as shown by a reduced reaction time to the same stimulus seen later on. However, in RSVP, there is no sign of priming for repeatedly seen distractor stimuli, even when they have been seen dozens of times.

Currently, the advent of fast graphics hardware has made it much easier to run RSVP experiments, and it is possible to present full screen images at up to 200 frames per second. At such speeds, the visual system is clearly overwhelmed. But remarkably, neurophysiological studies have shown that neurons at the top end of the monkey visual system can still respond selectively even when the images are changing at 72 frames per second. Specifically, they found that every time a neuron's favorite stimulus was shown, there would be a little "blip" of activation about 100 ms later—a delay that corresponds to the onset latency of the neuron. This is striking because it presumably means that several different images are simultaneously being processed at different levels in the visual system. While one image is still being processed in the retina, another is in the thalamus, another in the visual cortex, and so on. It recent years, a number of groups have used RSVP techniques to study the selectivity of neurons in the visual system, partly because by presenting as many as 10 images per second, it is possible to obtain a lot more data in a limited amount of time. This is of practical value because the time available to explore the properties of individual neurons is often limited, but it also allows researchers to examine the temporal constraints on visual processing.

Nevertheless, it is clear that at these very high rates both the strength of the neural response and the accompanying perceptual response are attenuated, compared to the situation where a single image is presented on its own. Typically, optimal activation would require a longer presentation time of 100 to 300 ms, a value that fits well with the typical fixation values seen with natural scene exploration. This is undoubtedly because the processing of any particular image will be affected by visual masking from both the preceding images (*forward masking*) and the images coming afterward (*backward masking*)—and the closer the images, the more masking there will be. Thus, while typical masking experiments will present a mask either before or after the target, RSVP protocols effectively combine both types of masking at the same time.

Simon J. Thorpe

See also Change Detection; Object Perception; Response Time; Speed of Processing in Sensory Systems; Visual Masking

Further Readings

Coltheart, V. (1999). *Fleeting memories: Cognition of brief visual stimuli.* Cambridge: MIT Press.

Keysers, C., & Perrett, D. I. (2002). Visual masking and RSVP reveal neural competition. *Trends in Cognitive Science, 6*(3), 120–125.

Ogmen, H., & Breitmeyer, B. G. (2005). *The first half second: The microgenesis and temporal dynamics of unconscious and conscious visual processes.* Cambridge: MIT Press.

Subramaniam, S., Biederman, I., & Madigan, S. (2000). Accurate identification but no priming and chance recognition memory for pictures in RSVP sequences. *Visual Cognition, 7*(4), 511–535.

REACHING AND GRASPING

Although sensory information is used to perceive the world, ultimately perceptions must result in actions to have an effect. Traditionally, more research and popular interest have been devoted to sensory perception than sensorimotor actions, but that has been changing with growing awareness of the sophistication of such actions and the underlying neural mechanisms. In addition, awareness has grown regarding how body parts and potential actions influence perception. The importance of sensory information for planning and controlling actions can be noted simply by looking at the amount of brain power devoted to such processing. Sensory information from all three spatial senses—vision, audition, and touch—is integrated in the posterior parietal cortex, which constitutes about 20% of the cerebral cortex. The parietal lobes send output to motor control areas in the frontal lobes and to subcortical structures and they exchange information with brain areas in the temporal lobes involved in sensory perception. Many brain regions are involved in controlling movements of the eyes, head, body, and particularly the hands.

Given that humans are an exceptionally dexterous species, almost all of our actions upon objects are performed with our hands, and our most common actions are reaching and grasping. Reaching refers to the movement of the arm and hand to contact an object, as when pushing an elevator button or brushing a fly away. Reaching is often accompanied by grasping to acquire an object, as when picking up a coffee cup. This act of prehension, or the seizing of an object with a body part, is performed not only by human hands, but also monkey feet and tails, elephant trunks, and giraffe tongues. The focus of this entry is on reaching and prehension in humans and other primates. Moreover, the emphasis is on actions directed toward visual targets, as is usually the case, though of course the actions can also be directed toward targets based on touch or sound.

Types of Reaches and Grasps

In everyday life, most reaches are accompanied by an action upon an object (such as pushing or grasping it); however, in the laboratory, scientists often study reaching alone by having subjects reach out and touch a target (sometimes called reaching, reaching-to-point, pointing, or aiming). In some cases, scientists will study related movements, such as pointing without reaching—directing the finger toward a target without extending the arm—or reaching in a virtual environment, such as using a joystick to move a cursor to a target.

Although there are many types of grasps possible with the human hand, research has typically focused on two types. *Power grasps* involving curling the fingers and thumb around a relatively large object, such as when grasping the handle of a hammer. *Precision grasps* involve using a single finger (typically the index finger) and thumb to carefully acquire a relatively small object, such as when picking up a nail. In addition, other grasp types are possible, including a *whole hand grasp* (as in picking up a softball), a *hook grip* (as in carrying a suitcase), or a *tripod grasp* (as in using three fingers to pick up a piece of a jigsaw puzzle). Once a grasp has acquired the target object, it may further be explored or manipulated. Right-handers typically perform grasps with the dominant right hand; however, left-handers are more likely to use both hands equally.

Reaching and Grasping Kinematics

Many studies of reaching and grasping record kinematic measures, such as the trajectory, timing, accuracy, and hand preshaping, and in some cases, the forces applied to the target object. These variables are recorded by placing sensors or markers on the arm and hand, and in some cases, having subjects grasp an apparatus that measures finger pressure and lift force. During reaching, the hand accelerates, reaches peak velocity approximately halfway through the movement, and then decelerates as it approaches the target. As the target is moved further away, the velocity increases, such that the total time to execute the reach remains approximately constant. Reaches show a tradeoff between speed and accuracy (sometimes called Fitts' law); that is, if high accuracy is required, the velocity of the reach is slowed. Accuracy of reach-to-point movements can be divided into two types of error. *Constant error* refers to a regular error in the average endpoint, such as when the subject consistently overshoots the target. *Variable error* refers to the degree of scatter of endpoints across multiple trials. For example, a dart player who consistently hits a location 5 centimeters (cm) above the bull's-eye would have a large constant error and small variable error, whereas another player whose throws landed randomly all over the board would have a small constant error and large variable error.

Many additional kinematic measures are used to quantify grasping. During grasping, visual information is used to preshape the hand opening, such that the hand opens wider than the object, reaching the maximum opening approximately 70% of the way through the movement, before closing upon the object. This maximum grip aperture typically scales with the size of the object to be grasped; for example, the hand opens wider before acquiring a tomato compared to a grape. As the hand contacts the object, the fingers are placed strategically to form a stable grasp. For example, in a precision grip, the finger and thumb will typically be placed on opposite sides of a line that passes through the center of mass. In cases where the grasped object is lifted, the fingers close upon the object with a certain grip force that takes into account the surface friction and fragility of the object (e.g., stronger for a wet, slippery rock than a dry, fragile egg). If the object begins to slip from the fingers, the grip force must be increased. An upward lift force is also applied to raise the object, taking into account the weight of the object (e.g., stronger for a rock than a ping-pong ball). The grip and lift forces must be coordinated; for instance, heavy objects require both forces to be higher than light objects.

Visually guided actions involve both movement planning and control. That is, first the target must be identified and the arm action planned; then the action can be adjusted based on feedback about the accuracy. Researchers sometimes try to dissociate these phases by limiting feedback, for example, by making the participant wear goggles that close once the action is initiated. These actions without feedback (open loop) are less accurate than those with feedback (closed loop). Specifically, the movements become slower and the hand may open wider upon approach to allow for a greater margin of error in the grasp. Similarly, removal of feedback from touch can also disrupt reach-to-grasp actions. For instance, if the fingertips are anesthetized, the hand again opens wider, and, once grasped, the object is more likely to slip between the fingers. When visual feedback is available, errors in reaching can be quickly corrected. For example, if the target jumps while the hand is in flight, the hand trajectory can be corrected within 1/10 of a second (even if the subject is unaware of the jump).

One prominent but debated theory suggests that the reaching and grasping components of a reach-to-grasp movement are separable. According to the theory, reach-to-grasp actions involve (1) a transport component (reach) to bring the arm and hand to the location of the target object; and (2) a grip component (grasp) to preshape the hand according to the object's size, shape, and orientation. The transport component involves proximal muscles in the shoulder and arm; whereas the grip component involves distal muscles in the fingers and hand. Moreover, the transport component is thought to be sensitive to object characteristics, such as location and distance, whereas the grip component is thought to be sensitive to object characteristics, such as size and shape. Although the theory of separate transport and grip components has been highly influential in the study of reach-to-grasp movements, the two components must be tightly choreographed and, in fact, some researchers argue they cannot be differentiated.

Brain Mechanisms Involved in Reaching and Grasping

Evidence from cognitive neuroscience suggests that different brain areas may be specialized for reaching and/or grasping. Two areas in the primate parietal cortex, the medial intraparietal (MIP) cortex and visual area V6A, are activated by reaching (and together may comprise the "parietal reach region"). In recent years, efforts have begun to develop brain-machine interfaces (or neural prosthetics) that would allow patients (for example with spinal cord injuries) to control prosthetic limbs using electrodes implanted into reach-selective areas of the brain. Some patients who have damage to the parietal cortex, presumably in the MIP cortex and area V6A, demonstrate optic ataxia, inaccurate reaching to the location of an object, particularly when the eyes are directed away from the object. Two other primate brain regions, the anterior intraparietal (AIP) area and the ventral premotor cortex (PMv), both respond more during grasping than reaching. In the macaque monkey, neurons within AIP (and nearby regions) fire vigorously when the monkey manipulates an object, such as when picking a raisin out of a hole. These neurons may code the visual properties of the object (such as shape and orientation), the motor actions performed, or some combination. If AIP or PMv are disrupted—by chemical injections, brain stimulation, or brain damage—preshaping of the hand prior to object contact is impaired. Human patients with damage to AIP show impaired grasping but intact reaching. Some "mirror neurons" within PMv and the inferior parietal lobe respond not only when a monkey performs a grasp, but also when it watches another individual perform a grasp.

Reaching and Grasping Over the Life Span

Reaching and grasping show different developmental time courses. Week-old infants begin to show increasing control of reaching using vision and, after several months, infants can even use vision of a moving object to predict where to intercept it. Newborns, even premature ones, show a grasp reflex whereby stimulation of the palm leads to closure of the hand with enough force to support the infant's weight. However, only after the age of nine months can infants use vision to orient their hands appropriately for the object's orientation and size. Precision grips begin to appear around the age of one year. Reaching and grasping are also affected in the elderly. Older adults have slower, less precise, and less continuous movements.

Relationship With Eye Movements

In everyday life, eye gaze and hand location are tightly coupled. When a person plans a reaching or grasping movement, the eyes typically lead the hand. That is, in experiments where the eyes are tracked during hand actions, the gaze lands at target objects and obstacles just before the hand approaches. In part this occurs because the eye is lightweight compared to the arm and thus can be moved more quickly. In addition, having the gaze directed toward the target may make it easier for the brain to compute the desired location of the hand based by transforming information about the location of the object on the retina, the direction of gaze, and the starting location of the hand. When people are forced to fixate upon a point while reaching or grasping at another location, accuracy drops considerably.

Relationship With Perception

The ability to use vision for reaching and grasping of an object can be dissociated from the ability to recognize the presence or identity of the object. Some patients with damage to the visual cortex or its inputs may show blindsight, a phenomenon where they are able to reach toward and even grasp objects of which they are unaware. Other patients with visual agnosia who cannot recognize object shape, size, or orientation may nevertheless be able to preshape and scale the hand to grasp the object. Even in control subjects without brain damage, perception and action can be dissociated. Although perception is reliably fooled by illusions, reaching and grasping may remain accurate. For example, in one illusion (the Ebbinghaus or Titchener circles illusion), a circle appears larger when surrounded by smaller circles and smaller when surrounded by larger circles; yet in both cases, the hand opens to the same size and is not tricked by the illusion. Taken together, these results suggest that while sensory information is essential for both perception and action, the two functions may rely on different computations.

Jody C. Culham

See also Action and Vision; Agnosia: Visual; Brain Imaging; Embodied Perception; Eye and Limb Tracking; Eye Movements and Action in Everyday Life; Haptics; Kinesthesia; Mirror Neurons; Optic Ataxia; Perceptual Development: Visually Guided Reaching; Tool Use; Visually Guided Actions

Further Readings

Castiello, U. (2005). The neuroscience of grasping. *Nature Reviews Neuroscience, 6*(9), 726–736.

Culham, J. C., Cavina-Pratesi, C., & Singhal, A. (2006). The role of parietal cortex in visuomotor control: What have we learned from neuroimaging? *Neuropsychologia, 44*(13), 2668–2684.

Georgopoulos, A. P. (1986). On reaching. *Annual Review of Neuroscience, 9,* 147–170.

Jeannerod, M. (1997). *The cognitive neuroscience of action.* Oxford, UK: Blackwell.

Jones, L. A., & Lederman, S. J. (2006). *Human hand function.* Oxford, UK: Oxford University Press.

MacKenzie, C. L., & Iberall, T. (1994). *The grasping hand.* Amsterdam: Elsevier.

Wilson, F. R. (1998). *The hand.* New York: Vintage Books.

READING

See Eye Movements and Reading

READING TYPOGRAPHY

The configuration of text rendered by printed, mechanical, or electronic means encompasses most of what we call typography. Typography may have an impact on legibility, the ease with which strings of letter forms can be decoded into words by a competent reader, and it may serve to supplement or support the semantic content of a text communication. Reading, of course, is an impressive perceptual and cognitive feat, and studying the effects on it of both experimental manipulation and "naturalistic" observation of typographic variables may aid our understanding of more fundamental aspects of object and word recognition, as well as of reading itself. This entry discusses aspects of legibility, economic considerations, and some misconceptions of *reading topography.*

Effect of Typographical Features on Readability

Typographic effects on legibility and typographic communications may operate at both small and large scales. At small scales, the forms making up individual letters (known as *glyphs*) have an impact on readers' ability to distinguish and identify individual letters and hence to recognize words or symbols. Stylistic variants of letter forms, such as italics and boldface, may indicate emphasis, whereas stylized font families may convey cultural messages (e.g., "old-English," "Broadway," or cursive-style fonts). At larger scales, grosser features of typography, such as margins, columnar organization, and interletter (*kerning*) and interline spacing (termed *leading*), may also affect readability. Such features may also occasionally convey semantic content. An example of this is when very wide spacing (letter or line) is used to communicate opulence or sophistication.

How typography impacts legibility and readability has important applications in signage design. On highways, where drivers may need to read signs from great distance and at high speeds, typography may have an impact on safety. In this setting, the most important criterion of legibility is the minimum visual (angular) size of letters needed (or inversely, the maximum viewing distance allowable) for good recognition of sign content. This criterion may be termed a *visual acuity* criterion; although in the case of acuity, it is the reader's functional ability that is assessed, whereas in legibility studies, a reader's functional ability, assumed to be constant, is used as a yardstick for assessing effectiveness of typographic manipulation.

In recent years, many researchers have sought instead to use a *reading speed* criterion to define legibility. Despite advances in computer-controlled presentation and experimental paradigms, reading speed is still difficult to use in legibility studies because of the high variability between individual text samples and in reading competencies of individual research subjects. To detect small effects despite these sources of noise, such studies often require many experimental trials.

In studies of font glyph legibility, visual acuity and reading speed criteria generally agree—seldom are letter forms of one font found to be more legible by one criterion and less legible by the other. Letter spacing, however, does seem to have differential effects on the two criteria. Generous spacing helps make visually small text more legible, but reading is generally at least as fast, and sometimes faster, for closer-spaced text when letter sizes are substantially higher than threshold size, possibly because the eye movements required to read are more compact.

Economic Considerations

Typography, of course, may have an economic impact, in that setting text with larger type, wider spacing, or wider margins results in more space (e.g., print pages, sign dimensions) required to print the same content. Enhanced legibility, which generally requires more space, is sometimes compromised for the sake of economy. The balance point of this trade-off may shift with changing demographics, as the number of older readers and readers with impaired vision increases.

Font Point Size and Legibility

Letter size is conventionally characterized by the vertical size of a font, in points. (A point is most often defined as 1/72.27 inch or 1/72 inch in digital publishing applications, but there are many other definitions as well, all reasonably close to one another.) Prior to computer typesetting, there was no commonly accepted way to measure exact point size of a font: A 12-point font was one which appeared to be the same size as other 12-point fonts. Nowadays, font point size usually refers to the vertical height required to set lines of text vertically abutting but not overlapping (i.e., with no leading, or extra space, added). Fonts on computers and the Web are characterized by size measures in addition to points, such as pixels, but they also use only a vertical measurement.

It is well known (and perhaps obvious) that increasing point size increases font legibility. However, characterizing font size by vertical size alone fails to capture the impact of horizontal letter and word size, which (within limits) also generally enhance legibility, whether achieved by increases in glyph width or in interletter or interword spacing.

One such typographic variable that is known to affect font legibility is fixed width (each glyph fits into a "box" of fixed width, with space added as necessary) versus proportional width (box width varies with glyph width). Fixed-width fonts are more legible at very small character sizes (relative to visual acuity), whereas proportional fonts are more legible at medium and larger character sizes. Close-spaced type is also less legible at small sizes than wide-spaced type, and there is evidence that the effect of font proportionality is due merely to differences in letter spacing.

Another variable with a strong impact on legibility that is not captured by point size is x-height, which generally typifies the size of most of the lowercase letters. Thus, fonts with large x-heights (relative to point size) tend to be more legible because the glyphs are, on average, larger.

Other variables known to impact legibility are letter stroke width (which also usually varies, even within glyphs), with thick strokes tending to enhance legibility provided gaps and holes (termed *counters*) remain salient; and glyph width (or letter aspect ratio) with wider letters generally being more legible—a form of horizontal magnification. Figure 1 illustrates some key typographic parameters.

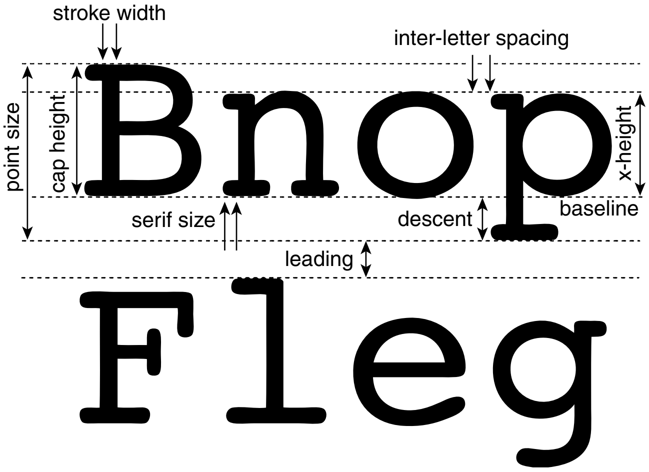


Figure 1 Some Key Typographic Parameters

Given variations in the letter shape and size within a single point size, it would seem that there is no simple way to characterize font size in a standard way that is meaningful across all fonts. Indeed, because some letters extend below and some above the baseline, characterizing even a single reference location for individual letters is a challenge.

Common Misconceptions

A couple of common misconceptions about font legibility are worth noting: First, many believe that serifs have a strong impact on legibility (some believing that they increase legibility, whereas others believe the opposite). However, experimental evidence suggests only a miniscule enhancement effect of legibility by the addition of serifs, and only at visually small sizes. This is likely a secondary effect due to the slight increase in interletter spacing that is required to accommodate the serifs.

Second, it is often claimed that mixed-case and/or lowercase text is more legible than all uppercase text, due to more distinctive letter and word shapes. Again, controlled experiments show that for both acuity and reading speed legibility criteria, all uppercase text is more legible than lower- and mixed-case, when set in equal point sizes that are small relative to the reader's size threshold. With larger point sizes, the uppercase advantage disappears.

Fonts in Research

Comparing and ranking legibility of specific fonts is fairly straightforward, using a visual acuity, reading speed, or some other criterion. However, conclusions drawn are limited to the specific fonts chosen. Another approach to studying legibility is to design fonts parametrically so that they vary only in a characteristic of interest, such as stroke width or serif size. In this case, conclusions can be more general, but some may argue that special purpose fonts designed for such studies are less natural looking and unlikely to be used in ordinary printed matter.

Aries Arditi

See also Eye Movements and Reading; Object Perception; Rapid Serial Visual Presentation; Visual Acuity; Word Recognition

Further Readings

Arditi, A., & Cho, J. (2005). Serifs and font legibility. *Vision Research, 45*(23), 2926–2933.

Arditi, A., & Cho, J. (2007). Letter case, legibility and low vision. *Vision Research, 47,* 2499–2505.

Arditi, A., Knoblauch, K., and Grunwald, I. (1990). Reading with fixed and variable character pitch. *Journal of the Optical Society of America A, 7,* 2011–2015.

Knuth, D. E. (1986). *The METAFONT book.* Reading, MA: Addison-Wesley.

Legge, G. E. (2007). *Psychophysics of reading in normal and low vision.* Mahwah, NJ: Lawrence Erlbaum.

Tinker, M. A. (1963). *Legibility of print.* Ames: Iowa State University Press.

RECEPTIVE FIELDS

A central assumption of sensory neurobiology is that the neural substrate of perception is the electrical activity of the sensory neurons activated by a given stimulus, that is, that understanding how sensory neurons respond to sensory stimuli will lead to an understanding of how *organisms* respond to sensory stimuli. But no less important than *which* stimuli are effective is *where* stimuli must be located to elicit neural responses. The *receptive field* of a sensory neuron is the region in the sensory periphery, for example a portion of the retina or of the body surface, within which stimuli can influence the electrical activity of that cell. This entry discusses the determination of the receptive field, the nonclassical receptive field, and characterization of receptive field properties.

The concept of the receptive field is central to sensory neurobiology in providing a description of the location at which sensory stimuli must be presented to a neuron to elicit responses. Although a number of investigators in the late 19th and early 20th centuries measured electrical potentials in the eye and optic nerve in response to visual stimuli, the modern concept of the receptive field begins in 1938 with H. K. Hartline, who was the first to isolate and record from single frog optic nerve fibers. He defined the receptive field of a retinal ganglion cell as the retinal area from which an increase in the frequency of action potentials could be elicited. In 1953, H. B. Barlow and S. W. Kuffler extended this definition to include all areas of the retina within which stimulation could either excite or inhibit the ganglion cell response. Around this time, V. B. Mountcastle began describing the response properties of single neurons in the somatosensory thalamus and the cortex. Thus, the receptive field came to mean the region over which one could influence the firing of a cell, not simply drive it.

There is a serial and hierarchical organization of receptive field properties. Each sensory modality is composed of multiple brain areas; as one proceeds from receptor to thalamus to the primary sensory cortex and beyond, receptive fields have progressively more complex stimulus requirements. For example, in the auditory system, peripheral neurons may respond well to pure tones, while some central neurons respond better to frequency-modulated sounds. In the visual and somatosensory systems, receptors' receptive fields can be essentially circular or oval regions of retina or skin (see Figure 1). By contrast, in the thalamus, visual and somatosensory receptive fields are still essentially circular, but with center-surround antagonism—onset of a stimulus in one skin or retinal region can elicit responses, while surrounding regions elicit inhibition (indicated in Figure 1 by plus and minus signs). Offset of stimuli in these regions produce the opposite responses. This antagonism of stimulus effects at different locations is one manifestation of the phenomenon called *lateral inhibition,* the interaction among sensory neurons described in the earliest receptive field studies by Hartline, Barlow, and Kuffler. The optimal stimulus is therefore not a spatially uniform extended one, but rather a discrete spot of light or contact with the body surface, with contrast between central and surrounding

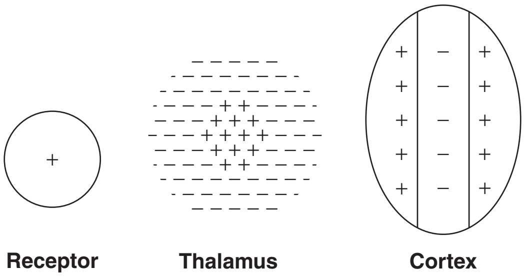


Figure 1 Profiles of Receptive Fields at Different Processing Stages

regions. In the primary visual and somatosensory cortex, receptive fields are further elaborated, being elongated and selective for the orientation and/or direction of motion of a stimulus. In higher visual cortical areas, neurons may respond best to images of faces or lizards or toilet brushes.

Referring as it does to a region, a receptive field is fundamentally a spatial entity (a portion of the visual field or retina, or a portion of the body surface); this makes the most sense in the visual and somatosensory systems. In the auditory system, as well, there is a spatial meaning to the term, as hair cells tuned to particular frequencies are located at different locations along the basilar membrane. Thus, in the auditory system one could define a cell's receptive field as the specific set of frequencies to which it responds. The receptive field of a sensory neuron anywhere in the nervous system is defined by its synaptic inputs; each cell's receptive field results from the combination of fields of all of the neurons providing input to it. Because inputs are not simply summed, referring to the receptive field properties of a neuron commonly means what stimuli the cell responds to.

Determination of the Receptive Field

While the receptive field may seem a simple concept, it has never been an entirely simple matter to define it, or how best to determine its extent. It is now appreciated that multiple definitions of the receptive field are required, and that the characteristics of a cell's receptive field (certainly its extent) depend on how it is measured. The classic method to determine the location and extent of the receptive field is to present discrete stimuli at different locations in the sensory periphery (on the retina, on the skin). The region that yields deviations in the spike discharge rate away from the background level has been variously referred to as the receptive field, the classical receptive field, the receptive field center, the discharge field or discharge center, the minimum discharge field, or the minimum response field. This includes the inhibitory subregions previously described (marked by minus signs in Figure 1) because stimuli presented in the inhibitory subregions can evoke responses when they are turned off. By this definition, stimuli presented outside of the cell's receptive field do not *by themselves* change its spiking activity.

The Nonclassical Receptive Field

However, stimuli presented in regions beyond the classical receptive field can still affect the cell; such influences can be either subthreshold or inhibitory (making them hard to detect if the cell has a low spontaneous activity). These influences can be revealed by pairing stimuli in the classical receptive field (which can include both excitatory and inhibitory subregions) with stimuli in the surrounding region. In the visual system, this has often been done by varying the size of sine wave grating stimuli centered over the classical receptive field and comparing responses to stimuli smaller than or exceeding that region. Typically, one finds that responses summate over a region greater than the classical receptive field; this is referred to as the summation area or summation field. Stimuli in the periphery of the receptive field may be unable to evoke responses by themselves, but can add their effect to responses driven by the more sensitive central portion of the receptive field. Summation fields could therefore simply reflect a spatial profile of sensitivity that is maximal at the center of the receptive field and that decreases to subthreshold levels at the periphery of the receptive field. The classical (minimum) receptive field is therefore a small portion of the region yielding responses; it is simply the most sensitive portion. In the visual cortex, the dimensions of this area are greater at lower stimulus contrast; thus there are high- and low-contrast summation fields. Coextensive with or larger than this summation region, there may also be inhibitory regions that do not themselves evoke responses. In different studies, the regions beyond the classical receptive field or summation field have been described as end zones, end-inhibitory zones, the silent surround, the nonclassical receptive field surround, the facilitatory or suppressive surround, or the modulatory surround. Modulatory surrounds are currently of considerable interest. A cell that responds with 30 impulses per second to a stimulus presented by itself in the classical receptive field can have its response doubled or abolished by pairing the same stimulus with another in the surround that by itself evokes no change in the cell's firing rate.

Thus, the receptive field first defined as the region within which sensory stimuli cause increases or decreases in firing can now be taken to mean the region either eliciting or modulating responses. One must attend to how the receptive field has been defined in each study. And one must bear in mind that there may be a classical receptive field surround (the region indicated by minus signs in the middle panel of Figure 1), as well as a nonclassical modulatory surround. One investigator's receptive field is another investigator's surround.

Characterization of Receptive Field Properties

There is another wrinkle to the understanding of receptive fields: a wrinkle in time. Several methods characterize not only the spatial structure of the receptive field but its temporal dynamics as well (the spatiotemporal receptive field), thereby capturing the fact that the spatial structure of the receptive field typically evolves over time, with excitatory and inhibitory subregions growing and shrinking in the period following presentation of a sensory stimulus. One approach, originally employed by G. L. Gerstein, is to define a peristimulus time (PST) response plane, in which response histograms are collected over time during and after stimulus presentation at a range of different locations. Spike-triggered averaging and reverse correlation (and other white-noise analysis) techniques are also employed to assess the spatial structure and stimulus selectivity of receptive fields and how these evolve over time. These techniques essentially look backward in time from the occurrence of a spike to determine what stimulus on average elicited that spike; in essence, this means computing a cross-correlation between the evoked spike train and the times and locations of stimulus occurrences. In the auditory system, this yields the spectrotemporal response field of auditory neurons.

Many studies have been devoted to characterizing not simply the extent and location of receptive fields, but also such characteristics as their shape, spatial organization, response timing, stimulus selectivity, or how their responses adapt. Describing the receptive field properties of a neuron means describing the stimulus selectivity of that neuron, for example, its selectivity for any stimulus parameter: size, loudness, velocity, color, spatial frequency, tone modulation frequency, and so on. It has been appreciated for some time that a number of receptive field properties (such as extent or spatial organization, i.e., degree of center-surround antagonism) can change with adaptation state. Thus, the apparent basic receptive field properties of a neuron are not entirely rigid; these depend on how they are measured, the adaptation state of the neuron, and the definition of a given property chosen by the investigator. The characterization of the receptive field properties of neurons informs us how single cells analyze the sensory world. The question remains: How are their collective responses put together to form sensory experience?

Jonathan B. Levitt

See also Audition; Cutaneous Perception; Neural Representation/Coding; Physiological Approach

Further Readings

Barlow, H. B. (1953). Summation and inhibition in the frog's retina. *Journal of Physiology (London), 119,* 69–88.

DiCarlo, J. J., & Johnson, K. O. (2002). Receptive field structure in cortical area 3b of the alert monkey. *Behavioural Brain Research, 135,* 167–178.

Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. *Current Opinion in Neurobiology, 10,* 438–443.

Gerstein, G. L., Butler, R. A., & Erulkar, S. D. (1968). Excitation and inhibition in cochlear nucleus. I. Tone-burst stimulation. *Journal of Neurophysiology, 31,* 526–536.

Hartline, H. K. (1938). The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *American Journal of Physiology, 121,* 400–415.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology (London), 160,* 106–154.

Iwamura, Y. (1998). Hierarchical somatosensory processing. *Current Opinion in Neurobiology, 8,* 522–528.

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology, 16,* 37–68.

Mountcastle, V. B., & Powell, T. P. S. (1959). Neural mechanisms subserving cutaneous sensibility, with special reference to the role of afferent inhibition in sensory perception and discrimination. *Bulletin of the Johns Hopkins Hospital, 105,* 201–232.

Ringach, D. L., & Shapley, R. (2004). Reverse correlation in neurophysiology. *Cognitive Science, 28,* 147–166.

Shamma, S. (2008). Characterizing auditory receptive fields. *Neuron, 58,* 829–831.

RECOGNITION

In a broad sense, the term *recognition* refers to the explicit feeling of familiarity that occurs when, for example, we view an object or hear a voice that we have experienced previously. It has been widely investigated in the visual domain, and this entry is thus based mainly on this field of research. By definition, accurate recognition can only occur for objects or sets of objects that we have experienced (seen) in the past. Recognition is fundamental to interpreting perceptual experiences, as it gives explicit meaning to our visual input.

Recognition presupposes two processing steps: (1) Previously seen objects have been stored in memory; (2) When an object is seen again, its image is compared to the stored representation. A match is successful when there is sufficient similarity between the image of the object and a stored representation.

There are also different levels of recognition, and our visual system can flexibly recognize objects at each of these different levels depending on the task requirements. For example, an animal can be recognized as belonging to a set of objects such as "a dog" (*categorization*) or as "my dog Bashi" (*identification*).

Problems Faced by Recognition Systems

We often do not realize the complexity of the processes underlying recognition. For example, when we locate (i.e., recognize) our displaced scissors on a cluttered desk, particular properties of the object, including illumination, orientation, and the distance from our eyes, may have changed from when we left it there. Furthermore, deformations of objects can also occur, such as scissors that are opened or closed. Consequently, the images that an object projects onto our retina can vary immensely. Despite such variations, our visual system is extremely effective at reliably matching currently viewed objects to the appropriate stored object representations.

Recent research has shown that humans are extremely quick at answering whether a flashed image contains an animal, for example; something they are able to accomplish within less than 400 milliseconds (ms). This implies that there is a direct route between the retinal projection of the object and the visual processing stages, with little time available for processing in feedback loops as the cortical response can be measured as early as 150 ms after stimulus onset. This is not to suggest that recognition always and only occurs without feedback; however, it indicates that there are strong time constraints on the recognition system.

There are certain types of recognition tasks that can be successfully accomplished by sophisticated artificial vision systems (e.g., face recognition systems used for security at banking machines). However, despite the fact that such technologies are rapidly progressing, the human perceptual system still remains unsurpassed in its ability to recognize arbitrary objects under many different environmental conditions and to adapt to new situations. For example, recognizing faces across changes in illumination and pose is a relatively trivial task for humans; however, it remains a largely unsolved problem for artificial systems.

Tasks Used to Evaluate Object Recognition

Numerous tasks have been devised to investigate the importance of factors such as illumination or depth information on recognition performance. Classical psychophysical approaches in the visual domain allow one to investigate the processes involved in recognition. For example, the importance of object orientation on recognition has been highlighted by research describing the face-inversion effect. Specifically, this is a phenomenon in which observers are faster at recognizing faces of familiar people shown upright than upside down. However, the limitation of purely psychophysical tasks is that demonstrating this effect does not tell us where in the brain orientation or face recognition processing occurs. Functional imaging techniques now allow researchers to extend their analysis to the level of the brain. Specifically, different activity patterns in the brain can be revealed, for example, by presenting observers with images of faces versus tools. Advancements in computer graphics and virtual reality technologies are also providing novel research opportunities. For example, investigators are now able to independently manipulate visual and haptic properties of objects as a way of investigating multisensory aspects of recognition.

There are a number of different *explicit* tasks, in which participants are required to evaluate objects as they are represented in the experiment. For example, in a *same-different task,* two images are shown to participants simultaneously or sequentially, and they are asked to judge whether these images are of the same object. In *match-to-sample tasks,* images of previously learned objects are shown among pictures of other, distracter objects. Participants must indicate which images match the previously learned object.

In *implicit* tasks, participants are not required to remember particular objects shown during the experiment. Instead, these tasks rely upon general knowledge of objects or object classes. In these tasks, for example, participants may be required to name objects as quickly as possible (*object naming task*), or classify faces as being male or female (*classification task*). In explicit and implicit tasks, viewing conditions of the objects, such as orientation or illumination, can be changed between images to test for the influence of such properties.

In most tasks, response times are faster for repeated objects or objects preceded by other objects related to them (e.g., fork followed by knife). Tasks specifically using this effect are called *priming tasks.*

In many recognition tasks, the stimuli must be presented quite rapidly. The rationale for this is to hinder the use of eye movements, or to increase task difficulty. In such cases where brief presentations are used, it is necessary to account for a phenomenon referred to as the *persistence of vision* (i.e., the fact that visual sensation is experienced longer than the physical presentation). To ensure precise presentation times, researchers often use a *mask.* A mask consists of another image, typically a random pattern, which is presented immediately after the test stimulus. This eliminates the persistence of the test stimulus as a way of avoiding residual perceptual processing.

Isabelle Bülthoff

See also Action and Vision; Computer Vision; Object Perception; Response Time; Visual System Structure

Further Readings

O'Toole, A. J., Phillips, P. J., Jiang, F., Ayyad, J., Pénard, N., & Abdi, H. (2007). Face recognition algorithms surpass humans matching faces across changes in illumination. *IEEE: Transactions on Pattern Analysis and Machine Intelligence, 29*(9), 1642–1646.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381,* 520–522.

Wallis, G., & Bülthoff, H. H. (1999). Learning to recognize objects. *Trends in Cognitive Sciences, 3,* 22–31.

RECOVERY OF VISION FOLLOWING BLINDNESS

In 1688, William Molyneux sent a letter to John Locke in which he asked whether a man born blind, who had learned to distinguish a globe and a cube by touch, would be able to distinguish them by vision alone, if sight were ever restored. It was not until almost 30 years after the posing of Molyneax's problem that the surgeon William Cheselden reported that a patient whose sight had been restored after years of blindness did indeed have acute difficulties interpreting the visual world. Trying to understand the extent and cause of these difficulties played an important role in the development of 18th-century philosophy as empiricist philosophers, such as Etienne Bonnot de Condillac, Denis Diderot, and Thomas Reid, refined their opinions about the relationship between perceptual concepts and sensory experience, and began to differentiate internal psychological events, such as sensations, perceptions, and cognitive constructs. This entry discusses cases of sight recovery and restored visual abilities.

Cases of Sight Recovery

Although the first recorded case of sight recovery was in 1020 A.D., only sporadic cases have been studied over the last 3 centuries. More notable cases include SB (1963), HS (1971), HB (1974), and more recently Virgil (1995), MM (2003), and SRD (2006). It should be noted that none of the cases of sight recovery currently recorded in the literature are considered "pure" sight recovery as defined by no light perception from birth to adulthood. Because all patients had intact retinas, they necessarily had some light perception, and in many cases rudimentary form vision, preoperatively. In most cases, sight recovery has been due to cataract removal in adulthood. Because of the current emphasis on early diagnosis and treatment of visual disorders, sight recovery in adulthood now generally occurs as a result of misdiagnosis, a lack of medical facilities, or associated ophthalmological complications that require treatment before successful cataract removal.

As well as differing in the amount of preoperative vision, most sight recovery patients were only studied some months or years after sight recovery had occurred. Nonetheless, despite these important differences across patients, some consensus has gradually emerged about the restored visual abilities of those who have lost their sight early in childhood.

Restored Visual Abilities

Most sight-recovery patients can name colors easily and can distinguish fine differences in hue. Motion processing also appears to be relatively spared. It was said of SB that the only visual objects he appreciated were those that were moving, such as birds. Similarly, for HB, it was reported that she could see the pigeons as they alighted in Trafalgar Square, but said that they appeared to vanish as they came to rest. MM had no difficulty on motion tasks that included identifying the direction of motion of bars, segregating textured fields based on motion, and using motion cues to compute the three-dimensional (3-D) shape of a rotating cube (see Figure 1a). MM was also sensitive to biological motion, recognizing a biological motion point-light "Johansson" figure, and was even able to resolve the fine cues that differentiate male and female gaits.

Sight recovery patients also have few difficulties with recognizing two-dimensional (2-D) shapes. MM could discriminate dots arranged in a circular structure (glass patterns) from dots arranged randomly, segment texture patterns based on luminance contrast, and identify whether a field of short lines contained a sequence of nearly collinear segments. The only 2-D task MM had difficulty with might perhaps be considered a 3-D task: Though he recognized outlined 2-D shapes, he could not identify the same shapes when defined by illusory contours (Kanizsa figures).

In contrast, sight recovery patients seem to have little understanding of 3-D shape or depth. According to Alberto Valvo, HS described his initial experiences after sight recovery as having no appreciation of depth or distance; he reported that street lights were like luminous stains stuck to window panes and the corridors of the hospital resembled black holes. MM could exploit occlusion cues but not shading, transparency, or perspective. He could not identify wire drawings of stationary 3-D shapes, describing a stationary cube as a square with lines, and was immune to illusions based on perspective cues, such as the Shepard tables (see Figure 1a). This difficulty in understanding 3-D percepts seems to result in great difficulties in recognizing even familiar objects and faces. Cheselden described his patient as being unable to visually discriminate a cat from a dog: Only by touching the animal was he able to recognize it as a cat. MM has described a similar reliance on touch to disambiguate visual information and shows deficits in object recognition tasks, as well as difficulties in recognizing the expression and gender of face images.

Why is it that color, 2-D shape, and motion perception are relatively unaffected by deprivation, while 3-D processing is dramatically impaired? One hypothesis is that some abilities might be more innate, develop earlier in infancy, and be more

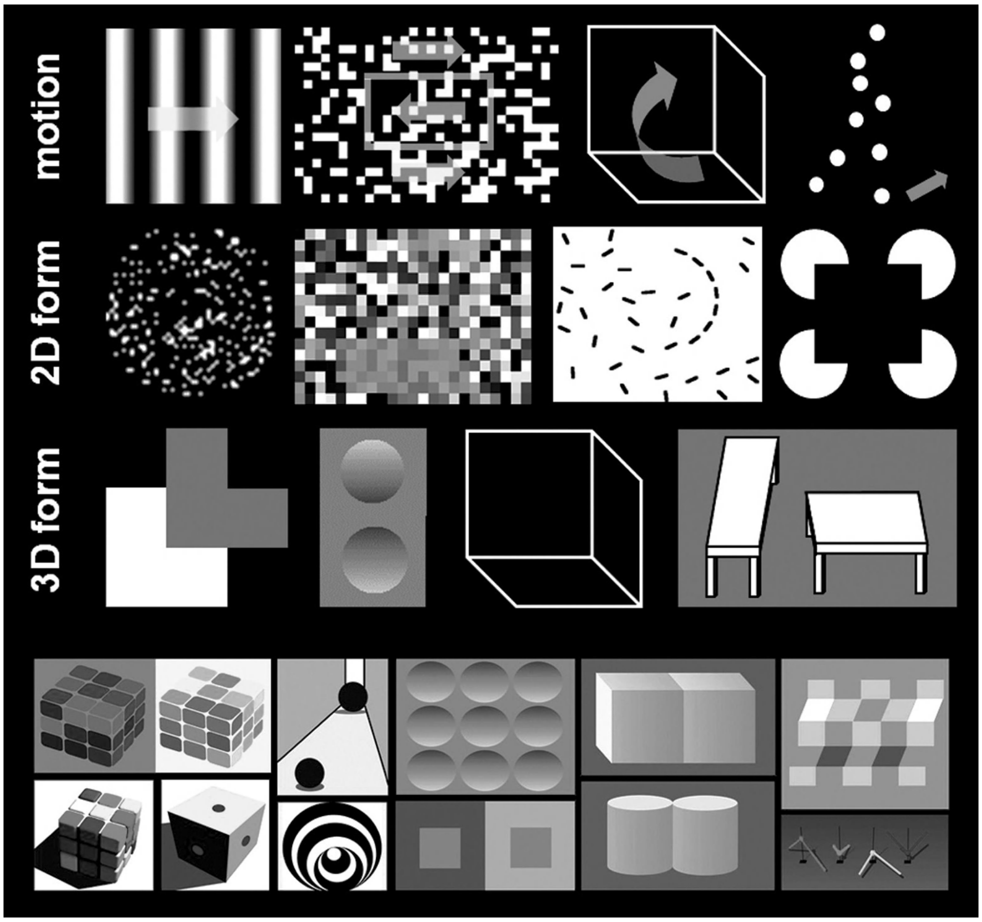


Figure 1 Examples of Motion, 2-D, and 3-D Stimuli Used to Examine Visual Processing in Sight Recovery Subject MM

robust to deprivation. However, the ability to perform some 2-D tasks, such as being able to trace the contours of objects, develops relatively late, yet can be performed by sight recovery patients. Nor do the pattern of deficits shown by sight recovery patients map neatly onto the division between ventral "what/recognition" pathways and dorsal "where/action" pathways. Although higher level object processing is severely impaired, there is a relative sparing of 2-D form processing and chromatic discrimination abilities. One intriguing possibility is that it is those visual abilities that depend upon an interaction between visual and tactile experience in development that are selectively disrupted. Perhaps sight recovery patients have no difficulties with purely visual sensations (color, 2-D shape, and motion) but, if the developmental experience of disambiguating visual experience with the help of touch is disrupted early enough in life, they fail to develop (or lose) the ability to construct a 3-D world from a 2-D retinal image.

It is still unclear whether visual performance can improve after sight recovery in adulthood. Some patients, such as HA and SRD, seem to have shown gradual improvement in the months and years after sight recovery. However, HA was not deprived of sight until the age of 15, and SRD recovered sight at the age of 12. Most other studies do not find evidence for postoperative improvement in visual performance, even in those patients that have been studied over considerable periods of time.

Ione Fine

See also Molyneaux's Question; Motion Perception: Physiological

Further Readings

Cheselden, W. (1728). An account of some observations made by a young gentleman, who was born blind, or who lost his sight so early, that he had no remembrance of ever having seen, and was couch'd between 13 and 14 years of age. *Philosophical Transactions, 402,* 447–450.

Fine, I., Wade, A., Boynton, G. M. B., Brewer, A., May, M., Wandell, B., et al. (2003). The neural and functional effects of long-term visual deprivation on human cortex. *Nature Neuroscience, 6,* 915–916.

Gregory, R. L., & Wallace, J. G. (1963). *Recovery from early blindness: A case study* (Experimental Psychological Society Monograph No. 2). Cambridge, UK: Heffer and Sons.

Kurson, R. (2007). *Crashing through: A true story of risk, adventure, and the man who dared to see.* New York: Random House.

Locke, J. (1690). *An essay concerning human understanding.* London: Eliz. Holt for Thomas Basset.

Ostrovsky, Y., Andalman, A., & Sinha, P. (2006). Vision following extended congenital blindness. *Psychological Science, 17*(12), 1009–1010.

Sacks, O. (1995). To see and not to see. *An anthropologist on Mars* (pp. 108–152). New York: Random House.

Valvo, A. (1971). *Sight restoration after long-term blindness: The problems and behavior patterns of visual rehabilitation.* New York: American Foundation for the Blind.

REORGANIZATION

See Cortical Reorganization Following Damage

RESPONSE TIME

In any task where a stimulus is presented and a response is emitted, the *response time* (RT) is the time between stimulus onset and response completion. In simple tasks with only one response that is given as quickly as possible no matter what the stimulus, RT is commonly referred to as *reaction time*. The methods and models researchers have used to study RT will be discussed in this entry.

It takes longer to make a difficult perceptual decision than to make one that is simple. Therefore, if we observe a person making a series of decisions, we can order those decisions by their difficulty by comparing their RTs. The use of RT to learn about mental events has a long history in psychology that dates back at least to the pioneering work of the Dutch physiologist F. C. Donders in the 1860s.

Donders developed a method of measuring the duration of different perceptual or cognitive stages called the *method of subtraction.* The idea was that by subtracting the RT of a simple task from the RT of a more complex task, one could estimate the time required to complete the extra processing stages required in the complex task. For example, in a task where either a blue or purple light appears on each trial, RT will generally be faster in a condition with only one response key that is depressed as soon as either light appears, than in a condition where the subject pushes one key if the light is blue and a different key if the light is purple. Donders argued that by subtracting the RTs in these two conditions, one could estimate the time it takes to discriminate the color blue from the color purple.

The method of subtraction is still popular today, and it has recently been extended to dependent measures other than RT. For example, a version of this method is widely used in functional neuroimaging experiments, where the neural activations of a simpler task are subtracted from those of a more complex task to determine which extra brain regions are required by the complex task.

The method of subtraction is based on a number of controversial assumptions. One is that the task under study requires *discrete processing stages,* or in other words, that an ensuing stage does not begin until exactly the moment at which the preceding stage completes processing. Neuroscience research suggests that many brain areas are simultaneously active in almost every task. This observation led to the development of *continuous-flow models,* in which an ensuing stage will begin processing before the preceding stage is completed.

A second assumption of the method of subtraction is serial processing stages, or in other words, that processing in multiple stages occurs sequentially. During the 1960s, parallel processing models were also developed. Parallel models assume that processing begins in all stages at the same time, although the finishing times of each stage may differ (e.g., as in a horse race). Parallel processing is strongly implicated in the pop-out effect, which occurs in certain visual search experiments. For example, when determining whether a red "T" is embedded in a background of black "T's," RT is virtually unaffected by the number of black

T's—that is, the red T seems to pop out of the background. On the other hand, when subjects are asked to determine whether a "T" is embedded in a background of "L's," RT increases sharply with the number of L's. Although this result is often interpreted as evidence of serial processing, parallel models can also easily account for such RT increases.

A third assumption of the method of subtraction is *pure insertion*—namely, that the perceptual, cognitive, and motor processes required in the two tasks are identical, except that the more complex task requires some additional processing stages. A weaker alternative, called *selective influence,* assumes instead that the two tasks require the same stages but in the second task the duration of one stage is different than in the first task. For example, printing the word *cat* in gray rather than black ink should increase the time needed to read the word, but it should not affect the subsequent time to decide that a cat is a mammal.

It has long been known that RT is closely related to response accuracy. For example, when people are instructed to increase their speed in a task, accuracy tends to decrease—a relationship known as the *speed-accuracy tradeoff.* The most popular of the earlier RT models and methods of data analysis ignored accuracy. Another line of research developed models that could simultaneously account for both RT and accuracy. The most popular of these modeled binary decisions as a random walk or diffusion process.

The idea behind the diffusion model is that as someone examines a stimulus that requires one of two possible responses, evidence accumulates gradually in favor of one response or the other. This accumulation is modeled as a diffusion process that increases when the evidence favors the first alternative and decreases when the evidence favors the second alternative. The first response is given at the moment when the accumulated evidence first exceeds some positive criterion and the second response is given when the evidence first falls below some negative criterion. As these two thresholds are moved further from zero, accuracy and RT both increase, so the diffusion model easily accounts for the speed–accuracy tradeoff.

Within the past few years, evidence has been reported that some cells in the brain that are associated with certain kinds of perceptual decisions seem to behave in a way that is consistent with the diffusion model. For example, cells in the lateral intraparietal cortex exhibit a push-pull pattern of responding that is consistent with the diffusion model in tasks when a monkey must decide whether to move its eyes to the right or the left.

F. Gregory Ashby

See also Decision Making, Perceptual; Rapid Serial Visual Presentation; Speed of Processing in Sensory Systems; Visual Search

Further Readings

Luce, R. D. (1986). *Response times.* Oxford, UK: Oxford University Press.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance* (Vol. 2). Amsterdam: North Holland.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes.* New York: Cambridge University Press.

RETINAL ANATOMY

Human vision operates over a huge range of light intensities, from the dimmest starlight to the brightest snow slope. Light enters the eye and is transformed by the retina into patterns of nerve impulses that pass up the optic nerve from eye to brain, and code the multiple dimensions of our vision: time, space, and color. The retina is a remarkable anatomical structure that achieves this transformation in a sophisticated manner at all these light levels, and optimizes the transmission of information to the brain through the fibers in the optic nerve.

The retina is a neural layer that grows out from the brain during early development. The retinas of all vertebrates share a similar structure. But from the perspective of human perception, it is the primate retina that is of most relevance, and is the focus of this entry.

Primate Retinal Basics

Among mammals, primates (monkeys, apes, and humans) are the sole group that possesses trichromatic color vision, that is, they possess three different types of receptors, called cones, which are responsible for daylight vision and code for color. They have peak sensitivities in the short (S, blue), middle (M, green) and long (L, red) regions of the spectrum. A further group of receptors, the rods, are responsible for vision in dim light, when we are colorblind. The visual systems of primates are very similar to ours, no matter whether anatomical, physiological, or behavioral criteria are used.

The retina lies at the back of the eye with the receptors farthest from the lens. The structure of the eye is sketched in Figure 29(a) of the color insert, with an enlargement of a small retinal circuit. After light has passed through the lens, it passes through the neural layers of the retina (which are virtually transparent) to be absorbed by the receptors. These are laid out in a regular array across the back of the eyeball, with the greatest density in the center of the visual field, which is called the *fovea*. The neural layers consist of bipolar cells that contact the cones and rods, and activate the ganglion cells, which send their axons and signals to the brain.

A cross section of the retina (color insert, Figure 29b) shows a layered structure with the receptors visible at the right as oblong, pencil-like elements pointing toward the incoming light. With the simple dye used to stain this section, it is impossible to disentangle the jumble of neuronal elements in the different layers. But with more sophisticated techniques, the layers and their neurons have been revealed and it is now possible to draw a wiring diagram of the retina, as seen in color insert, Figure 29(c). This looks complicated, but all pathways are variants of the same scheme—receptor to bipolar cell to ganglion cell.

At the right are the receptors; the cones are cone-shaped and the rods look like rods; the different types are color coded. The visual pigment that absorbs light and transforms it into electrical activity is situated in *lamellae*, which are stacked in the outer segments of the receptors. When light strikes the receptors and is absorbed, cones and rods hyperpolarize (their internal voltage goes from ~ –40 millivolts [mV] to more negative levels). Some of the different layers of the retina have been indicated: the receptor layer; the outer plexiform layer, where the bipolar cells contact the receptors; the inner plexiform layer, where the bipolars contact the ganglion cells; and the ganglion cell layer, which contains the ganglion cell bodies.

Retinal Pathways

Ewald Hering, a German scientist, proposed at the end of the 19th century that three separate channels exist in human vision, one responsible for black–white or achromatic vision, another responsible for the red–green dimension of our color experience, and a third responsible for the blue–yellow dimension. These are often called *opponent* channels. He believed this was so because it is possible to conceive of, for example, a reddish-yellow or greenish-blue but not a reddish-green or yellowish-blue. It now seems that such channels begin in the retina and visual pathways, segregating in the retina immediately after the receptors and staying separate up to the cortex. There is one more important concept; after the receptors, signals usually divide into two anatomical pathways, one excited by an increase in light intensity—the ON pathway—and another excited by a decrease in light intensity—the OFF pathway. There are typically both ON- and OFF-pathways for the different achromatic and color channels previously mentioned.

In color insert, Figure 29(a), the system responsible for achromatic vision is sketched at the top. After the cones, this channel begins in the diffuse bipolar cells, of which there are ON and OFF types, shaded light and dark in color insert, Figure 29(b). Both add the signals of L (red) and M (green) cones, and so are achromatic and noncolor selective. They feed into either ON or OFF parasol cells. These ganglion cells are so named because their dendritic trees (the branching part of the cell that receives input from the bipolar cells) are large and circular, like an umbrella. These cells send their axons to the magnocellular layers of the lateral geniculate nucleus, and are also called M or MC cells. One further point is that the dendritic trees of ON and OFF cells are carefully segregated anatomically, with the ON cells having branching dendrites in the inner half of the inner plexiform layer (near the ganglion cells), and the OFF cells having dendrites in the outer part of this layer. This can be seen in the wiring diagram (there are also differences in the way ON and OFF bipolars contact the receptors, which can be seen on careful inspection). As well as forming a correlate to an achromatic channel of vision, this system is the basis for the definition of *luminance,* which is a behaviorally defined international standard for light intensity used much in practical applications. It also appears to be the major input for motion perception.

The system responsible for red–green vision is unique to primates; there is no equivalent in other mammalian species. It can be seen in the middle paths of the wiring diagram. Single cones (L or M, red or green) contact just one ON or one OFF midget bipolar cell (as opposed to the diffuse bipolar cells that contact many cones), and each midget bipolar cell in the central retina contacts just a single midget ganglion cell. These cells are called "midgets" because they are so small, and have very small dendritic trees that contact just one other cone or cell. There are ON and OFF varieties of both of these cell groups. ON midget cells, which get excitatory input from just a single L (red) cone, are excited by red light and inhibited (i.e., switched off) by green light, while those that get input from a single M (green) cone are excited by green light and inhibited by red light. In both cases, the excitation comes from the cones directly, but the anatomical pathway by which the inhibition comes about is uncertain. OFF midget cells show the opposite pattern of responses; for example, a cell with inhibitory (OFF) input from an L cone via the midget bipolar is inhibited by red light and excited by green light. It might be thought that this cell type is the same as the ON midget that gets excitation from the M cone, but it appears that ON cells code for brighter colors and OFF cells for darker colors. Nevertheless, the dual nature of this pathway is symbolized by the red-green stripes used to label the midget cells and their associated cones. These midget cells send their axons to the parvocellular layers of the lateral geniculate nucleus and can be termed P or PC cells. They form the basis of red–green color vision and contribute to form vision.

The third, blue–yellow channel has its anatomical basis in at least two ganglion cell types. It is likely that an analogous channel is present in all mammals. There is an S-cone bipolar cell, shown as blue in the wiring diagram, which specifically contacts the S (blue) cones. The blue-ON ganglion cell is anatomically unusual: It has two layers of dendrites that laminate in the two levels of the inner plexiform layer, and because of this it is called a bistratified cell. One layer gets excited by S-cone ON bipolars, and the other gets inhibited by an OFF bipolar, which has not been identified. These cells are excited by blue light and inhibited by yellow light. Another cell type has been found that is excited by yellow light and inhibited by blue light, but its anatomical connections are uncertain. These two cells provide the yellow–blue dimension of our color perception.

These three groups of ganglion cells provide almost all of the input from the retina to the cerebral cortex via the lateral geniculate nucleus, and make up well over 90% of the ganglion cells in the retina, and are the foundation of our visual perception. There are other ganglion cell types that have been anatomically identified, but most send messages to the brain stem and are concerned with functions less to do with perception, such as setting pupil diameter.

Other Bits of Retinal Anatomy

There are other cell types in the retina. In the outer retina, there is a layer of cells in the outer plexiform layer, which make close contact with the cones and the ON bipolar cells. They are called horizontal cells and are also drawn into the wiring diagram. There are just two types in the primate, the so-called H1 cell, which adds signals from only the L (red) and M (green) cones, and the H2 cell, which, although it adds signals from all cones, is mostly driven by the S (blue) cones. These cells probably help the cones adapt to changes in light level. The inner plexiform layer contains a dense layer of contacts between ganglion and bipolar cells and another cell class called amacrine cells. There are many types (~30) of which the functions of just a few are known. As an example, one is drawn in the diagram. This is the AII amacrine, which gets input from ON rod bipolar cells. This is the only sort of rod bipolar, and an ingenious circuit generates on and off responses in ON and OFF ganglion cells respectively, that is, only one bipolar services both ON and OFF ganglion cells—mostly parasol cells. So it is with parasol cells that we see in starlight.

Retinal Philosophy

At first sight, retinal structure in primates appears largely determined by the trichromatic nature of vision in these species. There is good evidence that sensitivity to luminance, red–green and blue–yellow change is mediated by the different cell systems shown in color insert, Figure 29. But acuity and spatial vision are also critically dependent on retinal architecture, especially on the density and receptive field size of retinal ganglion cells. Retinal anatomy is similar from central vision up to the very edge of the visual field. It is just a matter of scale; the density of receptors, bipolar cells, and ganglion cells decreases drastically as one moves away from the fovea, and their receptive field sizes increase in a way that mirrors our loss in visual acuity. There is not full agreement on the relative roles of midget and parasol cells in fine detail vision, although the latter appear critical for the hyperacuities, our ability to localize objects with a precision far finer than the diameter of a photoreceptor. Parasol cells also provide the major input to brain areas responsible for motion perception. In conclusion, the retina is a remarkable part of the brain; an intricate network with precise connections that separates black–white and color signals for efficient information transfer up the optic nerve. It is one of the best understood of our sensory interfaces with the outside world but still holds a number of mysteries to be unraveled. For example, although the parasol, midget, and small bipolar cells appear in principle to correspond to the black–white, red–green, and blue–yellow perceptual channels proposed by Hering, there are important differences in detail and the reason for this mismatch is an unsolved problem.

Barry B. Lee

See also Animal Color Vision; Color Perception: Physiological; Light Measurement; Motion Perception: Physiological; Visual Processing: Retinal; Visual Processing: Subcortical Mechanisms for Gaze Control; Visual Receptors and Transduction

Further Readings

Boycott, B. B., & Dowling, J. E. (1969). Organization of the primate retina: Light microscopy. *Philosophical Transactions of the Royal Society of London, 255,* 109–184.

Dacey, D. M. (2000). Parallel pathways for spectral coding in the primate retina. *Annual Review of Neuroscience, 23,* 743–775.

Lee, B. B. (2004). Paths to color in the retina. *Clinical and Experimental Optometry, 87,* 239–248.

Wässle, H., & Boycott, B. B. (1991). Functional architecture of mammalian retina. *Physiological Reviews, 71,* 447–480.

REVERSE CORRELATION

Reverse correlation is a technique that allows a researcher to infer the collection of "features" used by an information processing system (e.g., human vision). Probably the best way to understand what reverse correlation is and how it works is through an example: Let's suppose you want to know what facial features a person uses to identify a human face. How might you go about solving this problem? A first guess might be to monitor an observer's eye movements as they perform a task that involves identifying human faces. However, this requires a strong assumption: Namely, that attention and the subsequent use of information are both perfectly correlated with eye movements. It is easy to show that this assumption is in fact not always true; for example, observers are able to shift their attention across different spatial locations while making minimal or no corresponding eye movements.

How then can we infer what parts of the visual scene an observer is using when performing a task? A less direct approach is to introduce random variations or "noise" to images shown to the observer and use this noise, along with the observer's responses, to work backward and determine the image locations where the noise had a significant impact on the observer's decisions. Such noise can be either introduced by the experimenter (as in our example) or naturally occurring within a system (such as the random variability that is part of the response properties of retinal ganglion cells). In our example of face perception, noise would refer to random values of brightness that an experimenter has added to the pixels of an image on a computer screen, much like the static you would see on a detuned television.

This kind of pixel-based noise can be added to every location in an image, and the observer can then be asked to perform a task with the noisy version of the image. To simplify matters, let's suppose we had just two faces that observers are asked to identify: one male face (Jacob) and one female face (Norma). On each trial of the experiment, the subject would see one of the two faces, chosen randomly with equal probability. In addition, a new sample of noise would be generated and added to each pixel in the image that the observer sees. If generated properly, the noise will make the task quite difficult, causing observers to make mistakes on about 30% of the trials.

Normally one might think achieving such poor performance in a task would be undesirable; however, in this case, having trials where observers make mistakes is highly informative. Consider a trial in which the observer is convinced the image is a noisy version of Jacob when in fact the image is a noisy version of Norma. Why would this occur? One possibility is that the observer has been influenced by their own "internal" noise, which has been shown to be present at several levels of visual processing. This might account for a small number of mistakes that the observer makes. However, more often, the observer has made a mistake because the noise has fallen on the image in such a way as to make it look slightly more like Norma than Jacob. That is, the noise that was added to the image has influenced the observer's decision process at some location or locations in the image and this has caused a classification error.

If we carry out this experiment on a computer, we can save the exact noise image that was shown on this trial as well as the observer's response and the true identity of the stimulus that was shown. If we then continue this experiment for many thousands of trials, we will end up with a large database of noise images and corresponding stimulus (S)–response (R) combinations. Suppose we sort the trials according to the different S-R combinations, so that there are four different categories of noise field: two that are associated with the response Norma (S = Jacob, R = Norma and S = Norma, R = Norma) and two that are associated with the response Jacob (S = Jacob, R = Jacob and S = Norma, R = Jacob). Next, suppose we compute an average noise field in each S-R category such that each pixel in an average noise image corresponds to the average value across all of the images at that particular pixel location. The resulting value in each pixel location the average noise images should be related to the relative influence of that pixel location on the observer's decisions across trials in the experiment. That is, those pixel locations where the noise had a big influence on the subject's responses should produce large values in the average noise images. In our face identification example, the average noise images for the two categories corresponding to the response Norma should reveal the collection of pixel locations that led the observer to make the classification choice Norma. Likewise, the remaining two average noise images should reveal the pixel locations that led the observer to classify the stimulus as Jacob.

In the domain of behavioral research, this approach (often called response classification rather than reverse correlation) has been used to measure the auditory and visual features used by human observers when performing a wide variety of perceptual tasks, such as grating and tone detection and discrimination, vernier acuity, motion discrimination, letter recognition, and object identification. Reverse correlation has also been used extensively to estimate the receptive field properties of sensory neurons by correlating neural activity with stimulus noise, rather than perceptual decisions. Thus, reverse correlation is an extremely powerful tool for specifying the processing that occurs in sensory and perceptual systems at both the levels of neurophysiology and behavior.

Jason M. Gold

See also Psychophysical Approach; Receptive Fields; Signal Detection Theory and Procedures

Further Readings

Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2002). Optimal methods for calculating classification images: Weighted sums. *Journal of Vision, 2*(1), 79–104.

Ringach, D., & Shapley, R. (2004). Reverse correlation in neurophysiology. *Cognitive Science, 28,* 147–166.

S

SCALING OF SENSORY MAGNITUDE

Imagine strolling along a boardwalk at noon on a sunny day; how brightly the sunlight makes the sea sparkle! It's time for lunch and you enter a dimly lit restaurant. At first you can't see a thing—all seems black! But in a few minutes, you can see well enough to read the menu. You have a nice lunch and then emerge, blinking, back into the bright sunlight. Again, you can't see a thing—only an all-pervading white glare! But again, after a few minutes, your vision is restored and you can view the sailboats prancing among the waves. How is it that your eyes can make these remarkable adjustments and provide clear vision under a huge range of light conditions? More specifically, and among other such questions, how does brightness (of, e.g., lamps in the dim restaurant, light reflected from the menu pages) increase with increasing time in the dark? This is a psychophysical question and the answer is given by a psychophysical scaling of the brightness sensation under experimental conditions. The resulting relationships between brightness of lights and time in the dark compose a set of fundamental facts that vision science must explain. They also constrain and inform studies of the physiological mechanism responsible for such dark adaptation. Any mechanism that purports to explain the experience of dark adaptation must predict these relationships. Many similar relationships among relevant variables for vision, hearing, taste, smell, touch, and pain experiences have been measured by psychophysical scaling, describing and elucidating phenomena such as adaptation and recovery, temporal and spatial summation, binocular and binaural summation, spatial inhibition, and so forth.

More generally, researchers measure the magnitude of various aspects of sensory or perceptual experiences in order to induce empirical laws, such as those of dark adaptation, as foundations of a complete understanding of human perception. Psychophysical scaling provides a useful and much-validated set of techniques to do this. A *scale* is a mathematical rule by which numbers are assigned to objects or events in order to measure some quantity associated with them. Modern measurement theory provides formal rationalizations for psychophysical scaling, but the techniques can be used without them. The minimum necessity is to understand that certain scale types allow a broader range of uses than do others, with *ratio* (ratios, differences, and zero are meaningful) and *interval* (only differences are meaningful) scales being the most useful. All techniques discussed yield interval scales at the minimum.

Some sensory or perceptual experiences can be arranged along *prothetic* continua (e.g., brightness, loudness, apparent size), whereas others cannot (e.g., shape, color, pitch). On prothetic continua, changes in stimulus intensity result in a change in the magnitude, the "how-muchness," of the perceptual experience. Prothetic continua are often represented in the brain by variations in the quantity of neural activity. Loudness, for example, seems to be represented by the total amount of neural activity in the auditory cortex. Prothetic continua can be meaningfully measured by all scaling techniques.

Other sensory or perceptual experiences can be ordered along a *metathetic* continuum that is qualitative rather than quantitative, involving changes in "kind." Here, changes in a stimulus (e.g., light wavelength) result in changes in a sensory quality (e.g., color). Some sense impressions have both prothetic and metathetic aspects; for example, light has both brightness and color. The techniques of this entry are not generally suited to measure metathetic continua, although sometimes they can be adapted by framing a quantitative question, such as "How similar do these colors appear to be?" or "How similar to the standard is this shape?" The similarity judgments can be scaled using nonmetric multidimensional scaling, and measurements of useful underlying quantitative variables can often be recovered.

As described in this entry, there are two basic approaches to psychophysical scaling. *Direct* scaling, in which observers assign a number directly to the magnitude of sensations, is the simplest and most straightforward. These techniques require the assumption that people can use numbers accurately according to their numerical value. *Indirect* scaling methods, in contrast, require the assumption that the quantity researchers wish to measure underlies people's ability to discriminate events that vary only in this quantity.

Indirect Scaling

In the 1850s, G. T. Fechner was the first to attempt to measure sensations; before that most thinkers believed it was impossible. Even Fechner believed that direct measurement was impossible. Nonetheless, he needed to measure sensations because he wished to establish a quantitative relationship between consciousness and its physical causes. He wanted to do this in the simplest way possible, and he believed that basic sensations, such as loudness, brightness, and so forth, were the least complex conscious experiences of which humans were capable. Thus, he chose to relate the magnitudes of conscious sensations to the magnitudes of the stimuli that cause them. He could already measure the smallest physical intensity difference between two stimuli that could just be noticed (the just noticeable difference, JND, or the difference threshold). In the 1830s, E. H. Weber had shown that the difference threshold increases as the reference intensity increases, now called Weber's law: $\Delta I = k \times I$, (i.e., the difference threshold, ΔI, is a fixed proportion, k, called the Weber fraction, of the reference stimulus intensity, I. However, Fechner thought that the *experience* of the difference between two sensations separated by one JND would be the same regardless of the physical size of the JND. Using calculus, Fechner reasoned that if Weber's law holds for a given sensory continuum, and all JNDs are subjectively equal, then $S = (1/k)\ln(I/I_0)$, where S is the magnitude of sensation a stimulus elicits (the number of JNDs of sensation above 0 at absolute threshold), I/I_0 is the physical magnitude of the stimulus (intensity, I, relative to the absolute threshold stimulus intensity, I_0), 1/k is the inverse of the Weber fraction, and ln means to take the logarithm to the base e of the following expression. This equation is called Fechner's law. It was the first statement of the now-accepted fact that the relationship between sensory experience and stimulus intensity is nonlinear. It was used for about 100 years to justify measuring sensation by a logarithmic transform of a physical intensity scale. For example, engineers often measure loudness using the decibels of sound intensity: $dB = 20\log(I/I_0)$, where $I_0 \approx 10^{-12} W/m^2$. The value of 1/k, which represents the unit of the scale, is different for different sensory continua because it is the inverse of the Weber fraction for the continuum scaled. For example, for electric shock $k \approx 0.01$ and $1/k \approx 100$, for light intensity $k \approx 0.08$ and $1/k \approx 12.5$, for sound intensity $k \approx 0.05$ and $1/ k \approx 2$, and for salt taste $k \approx 0.08$ and $1/k \approx 12.5$. Notice that if we multiply the unit of the Fechnarian scale for loudness of sound by 10 and transform to the base 10 logarithm, then we have decibels. Because Weber's law does hold true over most of the perceptible stimulus range for most sensory modalities, Fechner's law represents a way to

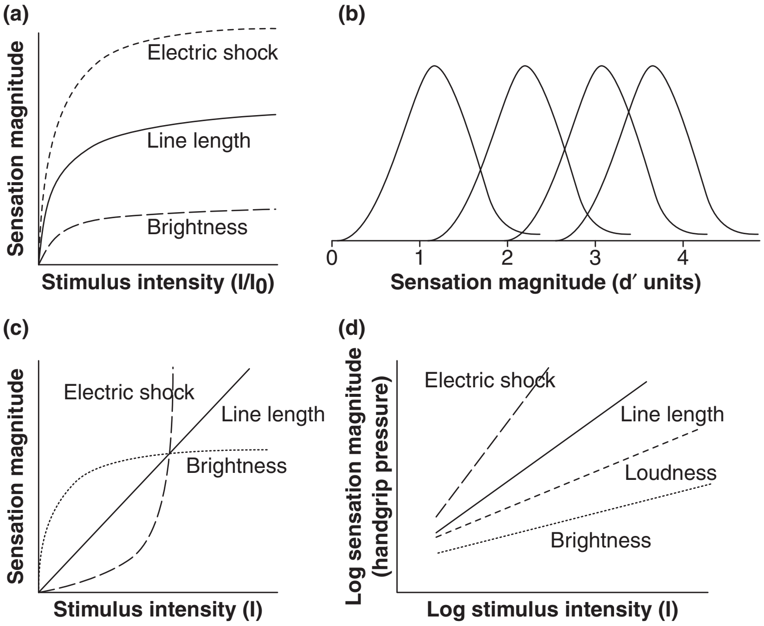


Figure 1 Scaling of Sensory Magnitude

Notes: (a) The psychophysical function relating stimulus intensity to sensation magnitude illustrated by representative Fechnarian scales for several continua. (b) Discriminability scales are based on inferred probability distributions of sensory response, the bell-shaped curves. (c) Stevens' law: power functions for brightness, length, and electric shock. Notice how different exponents give rise to different curves (relative stimulus intensity). (d) Cross-modality matching scales for four sensory continua with force of handgrip as the response continuum. Because the values on both axes are logarithmically spaced, all of the straight lines indicate power function relations between relative stimulus intensity and response magnitude. The line representing line length has an exponent of about 1.

obtain usable scales for every sensory or perceptual continuum for which a difference threshold can be measured. Figure 1(a) shows plots of sensation values against stimulus intensity for several modalities based on Fechner's law.

More modern indirect scaling techniques similarly involve measuring discriminability and adding up these measures to form a *discriminability scale*. In the 1920s, L. L. Thurstone devised a technique in which peoples' categorical (e.g., rate sensation magnitude from 1 to 10) or comparative judgments (which stimulus creates the greater sensation magnitude) were combined with several assumptions to yield estimates of the standard deviation of what he called *discriminal dispersions,* essentially hypothetical probability distributions over the response of a sensory system to repeated presentations of a particular stimulus (Figure 1b). As in Weber's law, these standard deviations typically increased with stimulus intensity. The scaling consisted of locating the means of the discriminal dispersions of a set of stimuli relative to an arbitrary zero point and spaced according to the number of standard deviations separating these means. In a variant of Thurstonian scaling, a discrimination measure such as d′ (the number of standard deviations of hypothetical sensory effect distributions separating a pair of not-perfectly discriminable stimuli), is computed between all adjacent pairs of stimulus intensity. A scale is formed by ordering the stimuli in terms of intensity relative to a threshold stimulus and then summing d′ distances between each successive pair to obtain a number for each stimulus in d′ units from the origin. Stimuli not in the tested set can be scaled by interpolation relative to the ones tested. Figure 1(b) shows a set of such stimuli with their hypothetical discriminal dispersions and scale values.

Direct Scaling

Direct scaling involves the direct assignment of numbers to sensations by the subject according to a rule stated by the experimenter. The subject must both understand the rule and be able to follow it. Both of these assumptions have been questioned, but the fact remains that direct scaling is both intuitive and useful.

Category Scaling

Bo Derek in the movie *10* was assigned the value 10 on the well-known 1-to-10 category scale for beauty. This is probably the most common and most used direct scale and, perhaps surprisingly, is useful scientifically as well. In category scaling, observers place the sensation created by each physical stimulus into 1 of N categories (typical values are N = 7, 9, or 10) from weakest to strongest. Stimuli are presented one at a time, and observers' average judgment for a particular stimulus intensity is treated as the scale value of sensation magnitude for that intensity. (Note that this is different from Thurstonian scaling, in which scale values are numbers of standard deviations from the origin.) Category scales are interval scales having an arbitrary zero but are used extensively nonetheless. It was at first thought that category scaling produced scales consistent with Fechner's law, but more recent studies show that category scales violate Fechner's law, as do all direct scales. In fact, category scales are consistent with the power law discussed next but with exponents about half the size.

Magnitude Estimation and Stevens' Law

In category scaling, available responses are limited to a few arbitrary category labels, usually numbers. Magnitude estimation asks observers to assign numbers to sensory experiences proportionally to the magnitude of those experiences and allows observers to use any positive numbers whatsoever, including decimals or fractions, to report sensation magnitude. Sometimes observers are given a number (modulus) that is to be assigned to a given sensory stimulus and others are judged relative to the modulus. In other applications, numbers are chosen freely by the observers. Again, stimuli are usually judged one at a time, and usually each is judged several times in one session. The average number assigned to each stimulus is then taken to be the scale value for that stimulus. Magnitude estimation gives rise to a ratio scale with a meaningful zero under certain conditions. It is still uncertain whether these conditions can be obtained in practice, but the scales are usually treated as ratio scales nonetheless.

S. S. Stevens had expected the results of magnitude estimation and other direct scaling experiments to confirm Fechner's law, but most did not. Over many experiments, Stevens found that, although the relationship was indeed nonlinear, as Fechner asserted, the equation that best described the relation of the average magnitude estimates of sensation to physical intensities was a power function, $S = aI^m$, where S is sensation intensity obtained from direct scaling, a is an arbitrary constant, and m is a characteristic exponent (power) that varies over sensory or perceptual continua. For some sensory continua, m is a small fraction (0.3 for brightness), for some it is close to 1 (line length), and for others it is substantially greater than 1 (up to 3.5 for electric shock, perhaps to allow us to avoid damaging intensities). Figure 1(c) shows that plots of power functions with different exponents (m) have dramatically different shapes, in contrast to those from Fechner's law. Exponents can be estimated from linear regression of log judgments on log stimulus intensities. This power function is often called the psychophysical power law or, after its popularizer, Stevens' law. As long as the experimental situation is kept reasonably standard and the same measures of physical stimulus intensity are used, the average exponents produced by different groups of observers for the same continuum are similar. Thus, once the standard value of m for a given sensory continuum is known, the scale value for any stimulus on that continuum can be computed from the power law. Using this principle, Stevens created standard scales recognized by the International Standards Organization (ISO) for several sensory continua; perhaps the most used is the sone scale for loudness: $S = 10.6\ P^{0.60}$, start here, where the physical measure, P, is sound pressure.

One problem with direct scaling is that results can vary more than is desirable across observers, experimental situations, or laboratories, because people use numbers in highly idiosyncratic ways. Solutions to this problem include transforming all data to a common master scale, attempting to find common anchor points to which conventional values can be assigned, and training all observers to a specific criterion of competence in use of a standard scale (constrained scaling). Each makes some progress toward eliminating the annoying variability inherent to direct scales, but currently none have been adopted as conventional.

Cross-Modality Matching

Magnitude estimation requires competence with number use, a competence that is not shared by all people, for example children, as well as a common conception of what is an acceptable way to use that competence. As previously mentioned, training alleviates some of the problems with number use but is counter to many researchers' feelings that direct scaling should not be constrained in this way. Stevens' solution to this problem was to invent a scaling procedure that does not involve numbers: cross-modality matching. In this procedure, an observer adjusts the intensity of a stimulus on one sensory continuum until the magnitude of the sensation it elicits seems to be equal to, or matches, that elicited by a stimulus from a different sensory continuum. Thus, an observer might be asked to squeeze a handgrip until the pressure feels as strong as a particular light is bright. Magnitude estimation is thus a type of cross-modality matching in which numbers are matched to sensations. Cross-modality matching data obey the power law for sensation intensities. Indeed, for canonical experimental situations and observers, magnitude estimates can be used to predict cross-modality matching data because both follow the power law; the cross-modality matching exponent is simply the ratio of the magnitude estimate exponents. For example, the cross-modality matching exponent for handgrip matched to loudness can be predicted from the exponents for handgrip and loudness scaled separately by magnitude estimation. Cross-modality matching scales are created in terms of a standard continuum, often as handgrip pressure as in Figure 1(d), although any continuum can form the standard, even numbers. Choice of standard continuum can be made as dictated by convenience in any particular experiment, although comparison across experiments would require a common standard continuum.

Lawrence M. Ward

See also Audition: Loudness; Psychophysical Approach; Psychophysics: Detection; Signal Detection Theory and Procedures

Further Readings

Bolanowski, S. J., Jr., & Gescheider, G. A. (Eds.). (1991). *Ratio scaling of psychological magnitude. In honor of the memory of S. S. Stevens.* Hillsdale, NJ: Erlbaum.

Marks, L. E. (1974). *Sensory processes: The new psychophysics.* New York: Academic Press.

Marks, L. E., & Gescheider, G. A. (2004). Psychophysical scaling. In H. Pashler & J. Wixted (Eds.), *Stevens' handbook of experimental psychology* (3rd ed., Vol. 4, pp. 91–138). New York: Wiley.

Poulton, E. C. (1989). *Bias in quantifying judgments.* Hillsdale, NJ: Erlbaum.

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural and social prospects.* New York: Wiley.

West, R. L., Ward, L. M., & Khosla, R. (2000). Beyond magnitude estimation: Constrained scaling and the elimination of idosyncratic response bias. *Perception & Psychophysics, 62,* 137–151.

SEEING AS

Seeing as refers to the experience of perceiving the same object in two or more distinct ways (for example, Jastrow's duck–rabbit picture, which looks like a duck when seen with a horizontal orientation and looks like a rabbit when seen with a vertical one). One might look at this figure and not see a duck in it or not see a rabbit in it. However, when one looks again, one sees the duck or rabbit that was not previously apparent. Alternative terms for *seeing as* are *aspect seeing* and *aspect perception.* As discussed in this entry, seeing as has connections with both perceiving and thinking, and therefore it links to a wide range of psychological phenomena.

The philosopher Ludwig Wittgenstein (1889–1951) devoted significant attention to seeing as. It featured importantly in part two of the *Philosophical Investigations* and both volumes of *Remarks on the Philosophy of Psychology* and *Last Writings on the Philosophy of Psychology.* He was influenced by Wolfgang Köhler's Gestalt psychology, and his discussion was primarily focused upon criticizing this. Köhler and Wittgenstein took a broadly similar stance in rejecting the view that what is immediately perceived is raw data from the senses, which we then make into objects of perception. They claimed there is immediate perception of defined and organized objects. For example, if four dots are appropriately placed at right angles to each other, they are not seen as dots but as the corners of a square. Wittgenstein investigated philosophical questions about the everyday employment of psychological concepts by describing how the words associated with these concepts are used. He was concerned with the variety and complexity of the usage of psychological words and what this showed about the distinctive features of everyday psychological concepts. Wittgenstein thought the concept of seeing as should be approached by examining the usage of the term *seeing as* and related expressions.

Differences in the usage of words associated with seeing as and seeing illuminate the similarities and differences between these concepts. For instance, Wittgenstein considered some uses of the word *see*. *See* can be employed in the sense of what is ordinarily perceived but also in cases of perceiving a likeness (such as facial resemblance) between two objects that was not previously evident. Wittgenstein thought the notion of seeing as denoted a collection of concepts about perception that were connected. The classic instance of seeing as is *aspect change*. Aspect change occurs when a previously unnoticed facet of an object becomes apparent, and thus what is perceived is seen differently. For example, consider a black Maltese cross on a white background inscribed in an octagon. One might see this as a black cross on a white background on one occasion and as a white cross on a black background on another. Similarly, in the duck–rabbit picture, aspect change would have occurred if one had moved from perceiving it as a duck to a rabbit or vice versa. Seeing as involves locating what is seen in a different context and includes finding new links and comparisons. This explains why altering the context of what is seen can change how it is seen. Aspect change is not an alteration of perception but of attitude. The notion that aspect change is an alteration of attitude is also evident from Wittgenstein's elucidation of seeing as via the notion of aspect blindness. *Aspect blindness* is the inability to experience aspect change and is a failing of imagination, not perception. There are various ways of understanding whether a person has experienced a change of aspect in a particular case. One of the ways this might be evident is that a person would talk about the same object in different ways, thereby indicating a change of aspect. For Wittgenstein, talk of this kind was a natural response to what is perceived. However, simply reproducing an object in a different way would not in itself indicate that a change of aspect had taken place.

The central issue in seeing as is accounting for what constitutes an alteration in what is perceived. The apparent puzzle that Wittgenstein sought to resolve was how one could see an object as different although it remained unchanged. This puzzle about aspect change raised the issue of whether seeing as should be regarded as an instance of seeing or thinking. The treatment of this matter dominated his discussion of the topic, but ultimately he thought that seeing as had affinities with both seeing and thinking. The recognition of these affinities explains how, in a case of seeing as, the perception remains the same while what is seen in the sense of thinking has changed (in the sense of there being a different interpretation). For Wittgenstein the importance of seeing as lay in the way it put problems about the concept of seeing into sharp focus. This was because the role that concepts play in perception was highlighted. He claimed that kinds of seeing as differed according to the extent of thought involved in them. There are instances of seeing as that cannot be expressed without the ability to use the appropriate concepts. For example, aspect change in the duck–rabbit case cannot be articulated without invoking the concepts of a duck and a rabbit. In contrast to this, expressing some kinds of seeing as does not require the use of concepts at all. For instance, in the Maltese cross, aspect change can be identified by tracing the lines that make up the perimeter of a cross. Wittgenstein thought that one cannot be in error about perceiving an aspect, and that such perception has a clearly identifiable start and finish.

Mark Addis

See also Aesthetic Appreciation of Pictures; Attention: Effect on Perception; Gestalt Approach; Philosophical Approaches; Pictorial Depiction and Perception

Further Readings

Budd, M. (1989). *Wittgenstein's philosophy of psychology*. London: Routledge.

Wittgenstein, L. (1980). *Remarks on the philosophy of psychology* (Vol. 1). Oxford: Basil Blackwell.

Wittgenstein, L. (1980). *Remarks on the philosophy of psychology* (Vol. 2). Oxford: Basil Blackwell.

Wittgenstein, L. (1984). *Philosophical investigations.* Oxford: Basil Blackwell.

SELECTIVE ADAPTATION

In perception, adaptation refers to the adjustment of sensitivity of sensory mechanisms in response to particular sensory inputs. For example, after putting on a wrist watch, the feeling of the watch touching the skin will soon be gone, a manifestation of sensory adaptation. Similarly, when wearing a pair of tinted glasses, initially the world will appear as tinted with the colors of the lens, but soon one will no longer notice the altered color and the world will look normally colored. As Aristotle noted more than 2,300 years ago, our senses are not static, and after looking at moving objects such as a waterfall, then static things could be seen as in motion. This is an example of seeing a motion aftereffect, which is another consequence of sensory adaptation. As the two examples show, sensory adaptation can make an observer less sensitive to constant input (e.g., wrist watch), and it can profoundly change a subsequent perception (e.g., waterfall). *Selective adaptation* refers to the selective reduction in sensitivity to the specific property observers are exposed to. This entry demonstrates the many ways selective adaptation influences perception.

A key feature of sensory adaptation is its selectivity. For example, after looking at a vertical set of closely spaced fine parallel lines (grating), one becomes less sensitive to gratings of a similar orientation and similar fineness, but sensitivity to gratings with very different orientations (e.g., horizontal) and spacing (e.g., coarse) would be much less affected. Selectivity of adaptation implies selectivity in neural coding or neural representation, which means that a selectively adapted input feature is coded or represented in the neural system independently from those unadapted input features. If sensory mechanisms can be considered as composed of multiple "channels" that are tuned to certain dimensions of the input information, then the degree of selectivity in adaption is simply a reflection of the overlap between these channels. The observation that visual pattern adaptation is orientation selective suggests that the human visual system contains multiple channels that are tuned to different orientations.

Adaptation as a Tool

The two features of selective adaptation—ubiquity and selectivity—make it a useful tool in the investigation of mechanisms of neural coding and representation. A demonstration of selective adaptation to a particular feature suggests that such a feature is represented in the neural system. For example, we can recognize faces from different viewing angles, but are there relatively independent representations of different views of a face in the visual neural system? It was shown that adapting to a face turned to one side (e.g., 30° to the left) would lead to a subsequently presented frontal face being seen as turned slightly the other way (e.g., 2° to the right); in addition, viewing a car turned sideways would not generate a viewpoint aftereffect upon viewing a frontal face. Such a selective face viewpoint aftereffect suggests the existence of neural mechanisms coding specific facial views.

In a classical demonstration of the power of selective adaptation in elucidating physiological mechanisms, Colin Blakemore and Fergus Campbell measured observers' contrast sensitivity function before and after adapting to a high-contrast grating. They showed that the largest contrast threshold elevation occurred when the adapting and test gratings had the same orientation (orientation selective); they also found that the contrast threshold elevation only occurred when test grating was within a relatively narrow bandwidth of the adapting spatial frequency. This experiment provided the strongest behavioral evidence for the existence of orientation and spatial frequency-selective neural channels in humans. Furthermore, the demonstration of selectively adapting a particular channel resulted in a dramatic change in perception and revealed a link between a particular neural response and perception—the orientation and spatial frequency selective adaptation experiment showed that the presumed neural channels that are tuned to different orientations and spatial frequencies are directly responsible for the perception of the corresponding patterns.

Selective adaptation also provides an effective way to study whether certain processes can occur in the absence of awareness or under conditions of inattention. For example, orientation-selective adaptation, a phenomenon occurring in the visual cortex, can be induced in the absence of awareness of an adapting orientation, suggesting that cortical neurons can represent orientation information without awareness. On the other hand, face-identity adaptation cannot occur without attending to an adapting face, suggesting face-identity encoding is dependent on attention.

Cross-Adaptation and Contingent Adaptation

Logically, when two features are initially processed in different channels but later converge, it is possible that adaptation could occur between them. This is called cross-adaptation (or transfer of adaptation). For example, adapting to a downward motion with the left eye alone would lead to an upward motion aftereffect with either the left or the right eye, even though the right eye was not exposed to the adapting motion. This is because signals entering the two eyes will converge at early cortical levels, and the motion aftereffect is primarily due to adaptation of direction-selective neurons in the cortex. Conversely, if a selective adaptation effect transferred across the two eyes, then it is a good indicator that the neural site of adaptation is cortical. Other examples of cross-adaptation include adapting to stimuli defined by one attribute (e.g., luminance) and testing with stimuli defined by another (e.g., disparity). Evidence of cross-attribute selective adaptation (e.g., tilt aftereffect or motion aftereffect) would support the existence of attribute-independent coding of the adapted property.

Selective adaptation to one feature can sometimes be made contingent on another feature. The best-known example of a contingent aftereffect is the McCollough effect, a color aftereffect contingent on orientation (see color insert, Figure 26). Many other contingent aftereffects have been demonstrated (e.g., contingency between color and motion or motion and disparity), though the interpretations of contingent aftereffects remain somewhat controversial.

Direct Measures of Neuronal Adaptation and Functional Interpretation of Adaptation

Selective adaptation has been studied at the neuronal level using single unit recording in animals, and more recently, in human cortical regions with noninvasive functional magnetic resonance imaging (fMRI). These studies provide strong and clear evidence of selective adaptation in neurons tuned to different properties, but at the same time, they also show that selective adaptation cannot be explained by a simple reduction of the overall sensitivity of the neurons tuned to the adapting stimulus, but are rather due to a change in the interactions in the local neural network. Selective adaptation must serve important functional purposes. The adaptation process allows the neural system to code and represent incoming information more efficiently and optimally by shifting the neuronal response range to better match that of the predominant sensory input. In short, the adaptation process "fits the mind to the world."

Selective Adaptation in Other Senses

Going back to Aristotle, as we did in the beginning of this entry, he also described adaptation in other sensory domains, such as that observers would have trouble hearing after hearing loud noises, and would have trouble smelling after smelling very strong odors. Although it is more commonly seen in vision, selective adaptation occurs in all sensory modalities. For example, the auditory system shows selective adaptation to specific frequencies of amplitude modulation, as well as the direction of a sweeping frequency. For example, in olfaction, selective adaptation to the strength of the odors could occur based on perceived odor similarity, as well as the odorants' structural similarity. Cross-adaptation between sensory modalities also occurs (e.g., adaptation to visual motion in depth induces a direction-specific loudness aftereffect in sound).

Sheng He

See also Aftereffects; Afterimages; Neural Representation/Coding; Olfactory Adaptation; Prism Adaptation; Taste Adaptation; Visual Light- and Dark-Adaptation

Further Readings

Barlow, H. (1990). A theory about the functional role and synaptic mechanism of visual aftereffects. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 363–375). Cambridge, UK: Cambridge University Press.

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology, 203*(1), 237–260.

Clifford, C. W. G., & Rhodes, G. (2005). *Fitting the mind to the world: Adaptation and after-effects in high-level vision.* Oxford, UK: Oxford University Press.

He, S., & MacLeod, D. (2001). Orientation-selective adaptation and tilt aftereffect from invisible patterns, *Nature, 411,* 473–476.

Movshon, J. A., & Lennie, P. (1979). Pattern-selective adaptation in visual cortical neurones. *Nature, 278*(5707), 850–852.

SELF-MOTION PERCEPTION

Self-motion perception refers to the ability to keep track of changes in one's position or orientation in space. For animals to move effectively in the world, they must be able to predict and detect the consequences of their actions. Some of the information that controls the perception of self-motion in humans is experienced quite vividly as self-velocity, but self-motion perception can also be experienced simply as a change in felt location. This entry concentrates on sources of information for self-motion, navigation, and the role of other sensory systems in self-motion.

Sources of Information for Self-Motion

Many distinct sensory and motor systems contribute to the human ability to keep track of the location of one's body in space. One is the vestibular system. Most people are familiar with the disruptions of controlled self-motion caused by being dizzy. Spinning in place for an extended period of time causes signals from the semicircular canals of the vestibular system to become adapted. Errant signals resulting from this adaptation can produce the vivid sensation of rotation, which may cause a person to lurch (and even fall down).

Normally, vestibular signals are experienced in concert with the characteristic visual self-motion signals. It was in the late 19th century that Ernst Mach posited the existence of a unified spatial sense based on his observations of *vection,* the vivid experience of self-motion produced by visual information alone. For example, the motion of an adjacent train can be enough to produce vection in a stationary observer; however, imagine being seated within a large cylinder with vertical stripes of black and white, if the cylinder begins to spin around you, you will soon find that you experience yourself as rotating within a stationary cylinder, even though it is you that is physically stationary. This is known as circular vection. Linear vection—a sense of moving forward in a straight line—can also be produced with appropriate visual stimulation. Linear vection is the basis for theme park "rides" in which one experiences forward self-motion based primarily on visual displays.

In addition to vestibular information and visual information, a third major source of information about self-motion comes from biomechanical activity itself. This has been demonstrated by the excellent performance of people asked to walk to a visually previewed target without further visual feedback (i.e., blindfolded). People have been shown to be quite accurate in somehow knowing where to stop for distances of dozens of meters, whereas performance at this task using vestibular signals alone (that is, when being passively transported while blindfolded) is not nearly so accurate. The updating of self-position seems automatic. For example, if stopped along the way to a previewed target (still blindfolded) and required to throw a beanbag the rest of the way—people do quite well. They can also walk blindfolded along a path off to one side of the target, and point accurately to where the target is because their sense of self-position is being updated.

Navigation

Normally locomotor self-motion is accompanied by visual and vestibular signals, and all these sources of information are integrated. Studies that have examined the relative contributions of visual and vestibular signals in navigation have suggested that vestibular signals are especially important to the accurate perception of changes of orientation, but that visual signals and biomechanical activity may dominate in the perception of distance traveled.

Path integration refers to a process by which we update our position in space by summing the vectors we have traveled (i.e., without external feedback). For example, knowing that you have walked 40 paces and then turned 90° to the right and walked 30 paces is sufficient to specify your starting location as being 50 paces away along a vector 127° to the right of your current heading. Further vectors can be added as well. Vestibular, visual, and locomotor sources of information are each sufficient to help update one's sense of position in space, and all three types of information are probably used. Estimation errors in path integration accumulate because each error is included in further integrations. However, in normal situations, human navigation does not depend exclusively on path integration because it is aided by visual layout and by landmarks that may be tied to cognitive maps. But well-calibrated path integration is an essential component of normal action with respect to the spatial environment.

Role of Other Sensory Systems

This entry has considered three major sources of information for the perception of self-motion, but any sensory system that bears on spatial location (including audition and touch) can play a role in self-motion perception. For example, when blindfolded, touch can stabilize one's position on a treadmill using a single point of manual contact with a railing. The auditory system can localize sources of sounds as a means of tracking one's own location relative to them. Auditory analysis of the way sound waves reflect off nearby surfaces can also provide information about spatial layout, as can sound "shadows" produced when walking past poles or trees between you and a busy roadway, for example. These sources of information are useful because self-motion perception can be based on any awareness of one's own change in position relative to a spatial feature of the environment. Self-motion perception is a flexible multimodal system.

Frank H. Durgin

See also Action and Vision; Multimodal Interactions: Spatial Perception In Touch and Vision; Navigation Through Spatial Layout; Spatial Layout Perception: Psychophysical; Vestibular System

Further Readings

Durgin, F. H., Pelah, A., Fox, L. F., Lewis, J. Y., Kane, R., & Walley, K. A. (2005). Self-motion perception during locomotor recalibration: More than meets the eye, *Journal of Experimental Psychology: Human Perception and Performance, 31,* 398–419.

Lappe, M., Jenkins, M., & Harris, L. R. (2007). Travel distance estimation from visual motion by leaky path integration. *Experimental Brain Research, 180,* 35–48.

Loomis, J. M., Da Silva, J. A., Fujita, N., & Fukusima, S. S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 906–921.

Mittelstaedt, M.-L., & Mittelstaedt, H. (2001). Idiothetic navigation in humans: Estimations of path length. *Experimental Brain Research, 139,* 318–332.

SENSE INTERACTIONS

See Multimodal Interactions: Color–Chemical; Multimodal Interactions: Neural Basis; Multimodal Interactions: Pain–Touch; Multimodal Interactions: Tactile–Auditory; Multimodal Interactions: Visual–Haptic

SENSORY CODING

See Neural Representation/Coding

SENSORY REHABILITATION

Sensory rehabilitation is the restoration of the ability to perform tasks normally requiring visual or auditory input, when such input is impaired or missing. The process usually involves a combination of training in new strategies and techniques and also harnessing assistive technology. This entry discusses rehabilitation methods for those with full or partial vision loss, as well as examples of auditory rehabilitation.

Blind Rehabilitation

Even if vision loss is complete, strategies and technologies using the remaining senses of hearing and touch can be successfully employed to facilitate reading, computer use, travel, and the vast majority of daily tasks. There is evidence that regions of the brain normally considered visual can be co-opted to process tactile and auditory input, as in the case of Braille reading and spatial localization.

Residential and day programs teach nonvisual skills for the many tasks involved in daily living, independent travel, and employment. Sighted people are often surprised to discover how many tasks can be accomplished without vision. Rehabilitation tools include methods of organizing and labeling items in the home or workplace to facilitate finding them, learning how to access talking books and newspapers, reading and writing Braille for written communication, and learning nonvisual skills for cooking and grooming.

A major skill taught in blind rehabilitation is that of independent travel (often referred to as "way-finding" or "orientation and mobility"). Specially trained guide dogs are used by some blind travelers, but most use a long white cane. Specific scanning patterns are taught in which the cane is tapped on the ground at the approximate location of the next footstep, generating clicking sounds that inform the user about the ground surface and result in echoes or reflections from the surrounding environment. Skilled cane users analyze these echoes to extract considerable information about the nature of the surroundings. Traffic, ambient sounds, and sound shadows (i.e., changes in sounds caused by objects between the source and the listener) add to this complex information, facilitating detection and analysis of environmental features such as buildings, rooms, hallways, and building entry ways. At stop lights, the sounds of parallel traffic stopping and starting are used to determine when to cross. The sound of receding traffic is used as an aiming point for straight travel. The direction of the sun on the face and many other subtle cues are harnessed to enable successful independent travel without vision.

Low Vision Rehabilitation

If vision loss is partial, abilities can often be improved by adopting new visual strategies. For example, blind spots (known as *scotomata*) in or near the center of the visual field due to macular disease make reading difficult, but by learning to fixate above, below, or to the left or right of the visual target, task performance can often be improved. Magnification is also often required (either optically, electronically, or through large print), because the resolution of retinal areas outside the central area or fovea is greatly reduced.

When the right or left half of the visual field is knocked out in both eyes due to a stroke, the condition is called a *hemianopia,* and can occur in combination with hemi-neglect (i.e., the individual is unaware of and "neglects" the right or left half of his or her space). Rehabilitation employs training strategies such as boundary marking to draw attention to the edge of a column of print or (for example, if the right field is invisible) learning to look at the last letter of each word rather than the first letter so that the entire word will be seen. In an effort to help detect objects located in the missing field area, mirrors or prisms may be used to shift part of the visual field sideways in one eye.

Standard acuity tests measure the ability to resolve targets of maximum contrast (black on white) under good lighting and underestimate the difficulty of seeing in the real world, where few targets offer such high contrast and lighting. Photographers know that the average reflectance of the world is 18%, so even a black object against such a background is much less distinguishable than print on a page. Most individuals with low vision have reduced ability to detect these low contrast targets and are also bothered by glare, especially outdoors.

Because of the previously mentioned effects, optimizing lighting and contrast is vital to successful low vision rehabilitation. Sometimes enough light can even make an apparent blind spot or scotoma disappear. Shielding the eyes from the glare of the sky, for example with a broad brimmed hat, is often helpful, as is the placement of light sources (e.g., over the shoulder) so they illuminate the task at hand but cannot shine directly into the eyes. Practical steps can also be taken to enhance contrast. A simple example is marking the edges of steps with contrasting paint or tape. Filling a white mug with coffee is easier than using a dark-colored one that does not contrast with the liquid. A light-colored countertop may be better for tasks involving dark objects and vice versa.

Auditory Rehabilitation

Auditory rehabilitation consists of more than the fitting of hearing aids or cochlear implants. In a manner analogous to visual rehabilitation, there are strategies for making the best use of remaining hearing. Practice, training, and acclimatization to the use of whatever aid is provided is part of the picture, but so is training in lipreading and speechreading, the use of clear speech by family members, and learning optimal positioning for best sound reception. Some telephones interfere with hearing aids and efforts to eliminate these problems are underway. Directional microphones and group listening devices are useful in specific listening environments. Special technical adaptations are available for specific vocational tasks, such as using a stethoscope.

John Brabyn

See also Ageing and Vision; Agnosia: Auditory; Assistive Technologies for the Blind; Audition: Disorders; Auditory Scene Analysis; Braille; Cochlear Implants: Controversy; Cochlear Implants: Technology; Contrast Perception; Echolocation; Hearing Aids; Low Vision; Neuropsychology of Perception; Prostheses: Visual; Sensory Restoration and Substitution; Speechreading; Tactile Map Reading; Visual Disorders: Blindness

Further Readings

Mormer, E., Ortmann, A., Palmer, C., & Seelman, K. (2007). Maximizing participation for people with hearing loss. In R. Cooper, H. Ohnabe, & D. Hobson (Eds.), *An introduction to rehabilitation engineering*. New York: Taylor and Francis.

Silverstone, B., Lang, M. A., Rosenthal, B., & Faye, E. E. (Eds.). (2000). *The lighthouse handbook on vision impairment and vision rehabilitation* (2 vols.). New York: Oxford University Press.

Stuen, C., Arditi, A., Horowitz, A., Lang, M. A., Rosenthal, B., & Seidman, K. (Eds.). (1999). *Vision rehabilitation: Assessment, intervention and outcomes*. Exton, PA: Swets and Zeitlinger.

SENSORY RESTORATION AND SUBSTITUTION

One way of coping with blindness and deafness has been some form of sensory substitution—allowing one or more remaining senses to take over for the nonfunctioning sense. Because the senses of vision, hearing, and touch all convey information about the physical and social environments, they can substitute for each other to varying extents. However, because unaided sensory substitution is only partially effective, humans have long improvised with artifices to facilitate sensory substitution (e.g., Braille and the long white cane for blindness, sign language for deafness, and the Tadoma method of speech reception used by deaf and blind people, in which the listener places the hand over the mouth and jaw of the speaker in order to sense articulatory speech information). In recent decades, the advent of electronics and computers has resulted in assistive technology that greatly enhances sensory substitution and, in some cases, allows for the restoration of a nonfunctioning sensory apparatus. This entry discusses the correction and restoration of the senses, as well as sensory substitutions by use of abstract meaning, amodal spatial representations, synesthesia, rote learning, and brain plasticity.

Sensory Correction and Restoration

Sensory correction is a way to ameliorate sensory loss prior to transduction, the stage at which light or sound is converted into neural activity. Optical correction, such as eyeglasses and contact lenses, and surgical correction have been employed over the years to correct for refractive errors in the optical media prior to the retina. For more serious deformations of the optical media, surgery has been used to provide a clear optical path. Likewise, hearing aids and surgery on the tiny bones of the middle ear have long been used to compensate for hearing losses short of profound deafness.

When blindness or deafness results from sensorineural loss in the retina or inner ear or from lesions in the central nervous system, the only recourse for people wishing replacement of the missing visual or auditory function is some form of sensory substitution or sensory restoration. As an example of the latter, one of the great success stories in modern times is the cochlear implant, a device that electrically stimulates auditory nerve fibers in place of transduction by the hair cells. Many people who were profoundly deaf are now able to understand ordinary speech, hear environmental sounds, and even enjoy music as the result of receiving cochlear implants. With blindness, sight restoration remains in the research stage because of enormous technical challenges. There are two primary approaches to remedying blindness: retinal and cortical prostheses. A retinal prosthesis involves electrically stimulating retinal neurons beyond the receptor layer with signals from a video camera. A cortical prosthesis involves direct stimulation of the visual cortex with input driven by a video camera.

Sensory Substitution

Sensory substitution offers several advantages over sensory replacement. First, sensory substitution is suitable even for patients suffering sensory loss because of lesions in the primary projection areas of the cortex. Second, because the interface with the substituting modality involves normal sensory stimulation, problems associated with implanting electrodes (such as infection) are avoided.

Assistive technology has been especially useful in ameliorating blindness. For example, accessibility to print and other forms of text has been vastly improved by way of electronic Braille displays and synthetic speech displays, including those driven by video cameras. For obstacle avoidance and sensing of the local environment, a number of ultrasonic sensors have been developed that use either auditory or tactile displays. For help with large-scale wayfinding, assistive technologies now include electronic signage, like the system of Talking Signs© and GPS-based navigation systems, both of which use auditory displays. For deaf people, improved access to spoken language has been made possible by automatic speech recognition coupled with a visual display of text.

Because the three spatial senses differ enormously in terms of sensory bandwidth (the amount of information each sense can convey to the brain about external stimuli and events), the hope that touch or hearing, aided by some single device (like a tactile display driven by a video camera), can serve as a wholesale replacement for vision is untenable. This being the case, sensory substitution needs to be directed toward enabling particular functions of everyday life (e.g., reading text, understanding speech, moving through the environment). These different functions vary tremendously in their informational requirements.

A principled approach to using one sense modality to substitute for another in connection with a particular function is to (a) identify what optical, acoustic, or other information is most effective in enabling that function and (b) determine how to transform this information into sensory signals that are effectively coupled to the substituting modality. The first step, then, requires research to identify what information is necessary to perform the function. Consider obstacle avoidance, for example. A person walking through a cluttered environment is able to avoid bumping into obstacles, usually by using vision with sufficient lighting. Precisely what visual information or other form of information (e.g., laser range finding) best affords obstacle avoidance? One way to address this question is to degrade a person's vision by limiting the field of view (e.g., by goggles) and limiting spatial resolution (e.g., by blurring) in order to determine what minimum values of field of view and resolution afford successful travel.

After the information needed for a given function has been identified, the next step is to couple the information to the substituting modality. This coupling involves two different issues, sensory bandwidth and the specificity of higher-level representation. The sensory bandwidth of a circumscribed region of the retina and associated visual pathway is determined by the field of view, the spatial resolution within the field of view, and the temporal resolution. Given the informational requirements of a particular function (e.g., obstacle avoidance), it must be determined whether the sensory bandwidth of the substituting modality is adequate to receive this information. No one has seriously considered using the tactile sense to substitute for vision in driving a car. In contrast, simpler functions, like using a tactile display coupled to a video camera to locate a glowing lamp in a darkened room, can be accomplished easily using tactile substitution of vision.

Even if the substituting modality has adequate sensory bandwidth to accommodate the source information, there is no guarante e that sensory substitution will be successfulbecause the higher-level processes of vision, hearing, and touch are highly specialized for the information that typically comes through those modalities. A good illustration is the difficulty of using vision to substitute for hearing. Even though vision has greater sensory bandwidth than hearing, currently there is no successful way of using vision to substitute for hearing in reception of the acoustic speech signal, evidently because speech makes use of brain processes that vision does not engage.

Based on what we know so far, we can identify some of the bases on which one sensory modality might substitute for another in enabling particular functions.

Abstract Meaning

One approach, which at this point is only a hope, is to develop computational algorithms for converting sensory signals into abstract representation ("meaning"), which would then serve for conversion into signals for any other sensory modality. Thus, video signals conveying the content of a complex visual scene might be converted into descriptions of the surfaces, objects, and intelligent agents within the scene, and these descriptions could be used to generate sequences of text or graphic symbols to be displayed using vision, hearing, or touch. Because research in artificial intelligence has yet to produce anything resembling such high-level descriptions in response to unconstrained sensory input, this basis for sensory substitution is unlikely to become a reality any time soon.

Amodal Spatial Representations

Because behavior takes place in a common three dimensional (3-D) physical space, coordination between the three spatial senses clearly occurs. This means that the sensory representations guiding action need to be commensurate to some extent. An intriguing possibility is that signals from the spatial senses give rise to amodal spatial representations in the brain that do not retain the modality-specific features of the input. For example, whether we learn the layout of furniture within a room using vision or touch, it is possible that the mental representation of layout used later in navigating within that space (either with eyes open or closed) retains neither the visual or tactile features experienced during learning.

Synesthesia

For rare individuals, synesthesia is a strong correlation between perceptual dimensions or features in one sensory modality with perceptual dimensions or features in another. For example, such an individual may imagine certain colors when hearing certain pitches. Because of its rarity, strong synesthesia cannot be the primary basis for sensory substitution. However, much milder forms occur in the larger population, indicating reliable associations between sensory modalities that can be exploited to produce more compatible mappings between the impaired and substituting modalities. For example, a device called the vOICe uses hearing to substitute for vision. Because of the natural correspondence between pitch and elevation in space (e.g., low-pitched tones are associated with lower elevation), the device uses the pitch of a pure tone to represent the vertical dimension of a graph or picture. The horizontal dimension of a graph or picture is represented by time or by directional variation from left to right in auditory space (e.g., using stereophonic sound). This device is successful for conveying simple two dimensional (2-D) patterns and isolated shapes and is claimed to be of value for some blind people in interpreting complex images and scenes.

Rote Learning

Even when there is neither the possibility of extracting meaning using artificial intelligence algorithms nor the possibility of mapping the source information in a natural way onto the substituting modality, effective sensory substitution is not completely ruled out. Because human beings, especially at a young age, enjoy great neural plasticity and a huge capacity for learning, there is always the possibility that they can learn mappings between two sensory modalities that differ in their higher-level interpretative mechanisms. As previously mentioned, the vOICe converts 2-D spatial images into time-varying auditory signals. Although the device is based on the natural correspondence between pitch and height in a 2-D figure, it seems unlikely that the higher-level interpretive mechanisms of hearing are suited to handle complex 2-D spatial images usually associated with vision. Still, it is possible that if such a device were used by a blind person from early in life, the person might eventually develop the ability to interpret auditory information from unconstrained natural scenes.

Brain Plasticity

In connection with his seminal work with the tactile vision substitution system, which turned achromatic images of objects from a video camera into patterns of vibration presented to a person's skin surface, Paul Bach-y-Rita speculated that substitution of vision by touch involved a reorganization of the brain, whereby the incoming somatosensory input came to be linked to and analyzed by visual cortical areas. Though a radical idea at the time, it has recently received confirmation by a variety of studies involving brain imaging and transcranial magnetic stimulation (TMS). For example, brain imaging research has shown that the primary visual cortex is activated by Braille reading in blind people and that temporary functional lesions of the visual cortex produced by TMS interfere with Braille reading. This remapping of tactile signals into the visual cortex seems to bode well for more successful sensory substitution devices in the future. Tempering this hope is the reality that the sensory bandwidth of tactile processing is so much lower than that of vision.

Jack M. Loomis

See also Braille; Cutaneous Perception; Guidance Systems for Blind People; Multimodal Interactions: Visual–Haptic; Synesthesia; Visual Disorders: Blindness

Further Readings

Bach-y-Rita, P. (1972). *Brain mechanisms in sensory substitution.* New York: Academic Press.

Cohen, L. G., Celnik, P., Pascual-Leone, A., Corwell, B., Faiz, L., Dambrosia, J., et al. (1997). Functional relevance of cross-modal plasticity in blind humans. *Nature, 389,* 180–183.

Harrison, J., & Baron-Cohen, S. (1997). Synaesthesia: An introduction. In S. Baron-Cohen & J. E. Harrison (Eds.), *Synaesthesia: Classic and contemporary readings.* Malden, MA: Blackwell.

Loomis, J. M., & Klatzky, R. L. (2007). Functional equivalence of spatial representations from vision, touch, and hearing: Relevance for sensory substitution. In J. J. Rieser, D. H. Ashmead, F. F. Ebner, & A. L. Corn (Eds.), *Blindness and brain plasticity in navigation and object perception* (pp. 155–184). New York: Lawrence Erlbaum.

Meijer, P. B. L. (1992). An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering, 39,* 112–121.

Niparko, J. K. (2000). *Cochlear implants: Principles and practices.* Philadelphia: Lippincott, Williams and Wilkins.

Normann, R. A. (1995). Visual neuroprosthetics: Functional vision for the blind. *IEEE Engineering in Medicine and Biology Magazine, 14*(1), 77–83.

Sadato, N., Pascual-Leone, A., Grafman, J., Ibanez, V., Deiber, M.-P., Dold, G., et al. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature, 380,* 526–528.

Warren, D. H., & Strelow, E. R. (1985). *Electronic spatial sensing for the blind.* Boston: Martinus Nijhoff.

SHAPE CONSTANCY

See Constancy

SIGNAL DETECTION THEORY AND PROCEDURES

Nearly all reasoning and decision making takes place in the presence of some uncertainty. *Signal detection theory* provides a precise language and graphic notation for analyzing decision making in the presence of uncertainty and also shows procedures for interpreting occurrence, criterion, and possible outcomes. For example, imagine that a radiologist examines a computerized tomography (CT) scan, looking for evidence of a lung tumor. Interpreting CT images is difficult and requires a lot of training. There is some uncertainty as to what is there or not. Is that white patch a tumor? If the radiologist offers the opinion that there is no tumor and is wrong, there could be grave consequences for the patient. Then again, if the radiologist incorrectly indicates there is a tumor, the patient will undergo unnecessary further cost, testing, and stress. As described in this entry, signal detection theory provides a model of this sort of decision task, and has direct application to experiments in perception, but it also offers a way to analyze many different kinds of decision problems.

Information and Criterion

In the previous example, there are two possibilities for the patient: either the "signal" (tumor in this case) is present or absent. There are also two possible decisions: either the decision maker (the radiologist in this case) thinks he or she sees the signal (the radiologist responds "yes") or does not (the radiologist responds "no"). Thus, there are four possible outcomes. Two of these are correct: a *correct reject* (the tumor is absent and the radiologist says "no") or a *hit* (the tumor is present and the radiologist says "yes"). The other two are incorrect responses: a *miss* (the tumor is present and the radiologist says "no") or a *false alarm* (the tumor is absent yet the radiologist says "yes").

The radiologist bases his or her decision on information. For example, healthy lungs have a characteristic shape. The presence of a tumor might distort that shape. Tumors may have different image characteristics: brighter or darker, or a different texture. With proper training, a doctor learns what to look for; with more practice and training, the doctor will be able to acquire more (and more reliable) information. Running another test (e.g., magnetic resonance imaging, MRI) can also be used to acquire more information. The effect of acquiring more information is to increase the likelihood of correct outcomes (a hit or correct rejection), while reducing the likelihood of errors (a false alarm or miss).

In addition to relying on information from medical tests, the medical profession encourages doctors to use their judgment. Different types of errors are not always equal. The doctor may feel that missing an opportunity for early diagnosis may mean the difference between life and death. A false alarm, on the other hand, may result only in a routine biopsy operation. Consequently, the doctor may choose to err toward "yes" (tumor present) decisions. However, other doctors under the same circumstances may feel that unnecessary surgeries (even routine ones) are bad (expensive, stressful, etc.), so they may chose to be more conservative and say "no" (no tumor) more often. They will miss more tumors, but they will be doing their part to reduce unnecessary surgeries. And they may feel that a tumor, if there really is one, will be picked up at the next checkup. These arguments are *not about information.* Two doctors, with equally good training, looking at the same CT scan, will have the same information. But they may utilize a different *criterion.* Indeed, the same doctor might use a different criterion for different patients. For example, the criterion might be shifted toward "no" responses for a patient with a higher risk of complications from a biopsy procedure. On the other hand, the criterion might be shifted toward "yes" responses for a patient with a family history of lung cancer.

Internal Response and Internal Noise

Any measurement or test result has inherent variability; repeating the measurement will give a slightly different result. Scientists call this *measurement noise.* If a doctor acquires a large number of CT scans of a patient's lungs, one after the next, each of them will be slightly different. There are many possible sources of noise. Perhaps the patient moves or breathes differently from one scan to the next, or perhaps the patient is positioned slightly differently. Every effort is made to reduce the noise (e.g., asking patients to hold still and hold their breath), but there is no way to completely eliminate it.

When people make decisions, there is another type of uncertainty called *internal noise.* To illustrate the idea of internal noise, suppose that the doctor has a set of tumor detector neurons in his or her brain; these neurons receive visual information from the doctor's eyes and the doctor monitors the responses of these neurons. These hypothetical tumor-detector neurons will give noisy and variable responses. Neurons signal information with action potentials (also called spikes), which are small, brief changes in electrical voltage that propagate along nerve fibers. The response of a neuron is typically quantified as the number of spikes per second. After one glance at a scan of a healthy lung, a hypothetical tumor detector neuron might fire 10 spikes per second. After a different glance at the identical scan, this neuron might fire 40 spikes per second.

It is not known precisely which neurons in the doctor's brain are used to perform this task, but there is some internal state, reflected by neural activity in the brain, that determines the doctor's impression about whether or not a tumor is present. This neural activity might be concentrated in just a few neurons or it might be distributed across a large number of neurons (the average response, the sum of all responses, a difference or ratio of responses, etc.). This hypothetical brain activity is called the decision maker's *internal response* (or *decision variable*). The internal response is inherently noisy. Even when there is no signal (no tumor in this case), there will be an internal response (sometimes more, sometimes less).

Probability of Occurrence, Criterion, and Possible Outcomes

Probability is the mathematical tool for describing variability of internal responses. Figure 1 shows a graph of two hypothetical internal response curves. The curve on the left is for the *noise alone* (no tumor), and the curve on the right is for the *signal plus noise* (tumor present). The horizontal axis is labeled internal response and the vertical axis is labeled probability. The height of each curve represents the likelihood that level of internal response will occur. To be concrete, the horizontal axis could be labeled in units of firing rate (10, 20, 30, etc., spikes per second). This would mean that for noise alone (no tumor), it is most likely that the internal response would be 10 spikes per second. It is also rather likely that the internal response would be 5 or 15 spikes per second, but very unlikely that the internal response would be 25 spikes per second when no tumor is present. To remain noncommittal about what and where in the brain the internal response is, the horizontal axis has not been labeled in terms of firing rates. The internal response is in some unknown, but quantifiable, units.

The decision maker must interpret these neural signals, the internal response, and then respond "yes" or "no." The simplest strategy that the decision maker can adopt is to pick a criterion along the internal response axis, responding "yes" whenever the internal response is greater than this criterion and "no" whenever the internal response is less than this criterion. One possible criterion is indicated by the vertical line in Figure 1. The criterion line divides the graph into four sections that correspond to hits, misses, false alarms, and correct rejections. On both hits and false alarms, the internal response is greater than the criterion, so that the decision maker responds "yes" to these internal responses. Hits, for example, correspond to signal plus noise trials when the internal response is greater than the criterion; this is indicated in the figure as the area under the signal plus noise curve to the right of the criterion.

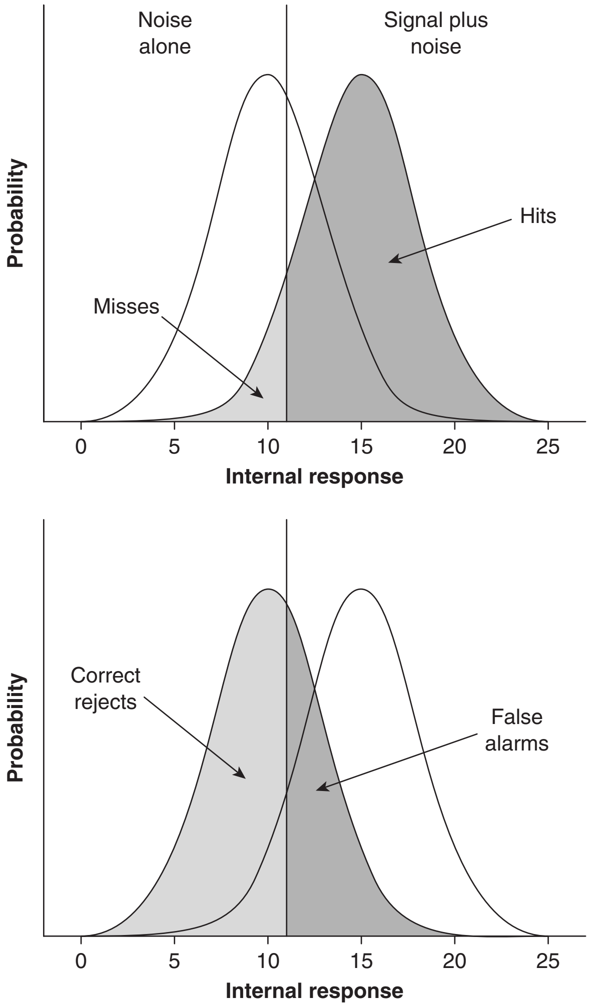


Figure 1 Internal Response, Criterion, and Possible Outcomes

Notes: Noise alone distribution (curve on left) and signal plus noise distribution (curve on right). The criterion is indicated by the vertical line. The decision maker says *yes* when the internal signal is to the right of the criterion. In the top panel, the signal is present (shaded area)—so saying *yes* above the criterion is a hit and saying *no* below the criterion is a miss. In the bottom panel, no signal is presented (shaded area)—so saying *yes* above the criterion is a false alarm and saying *no* below the criterion is a correct rejection.

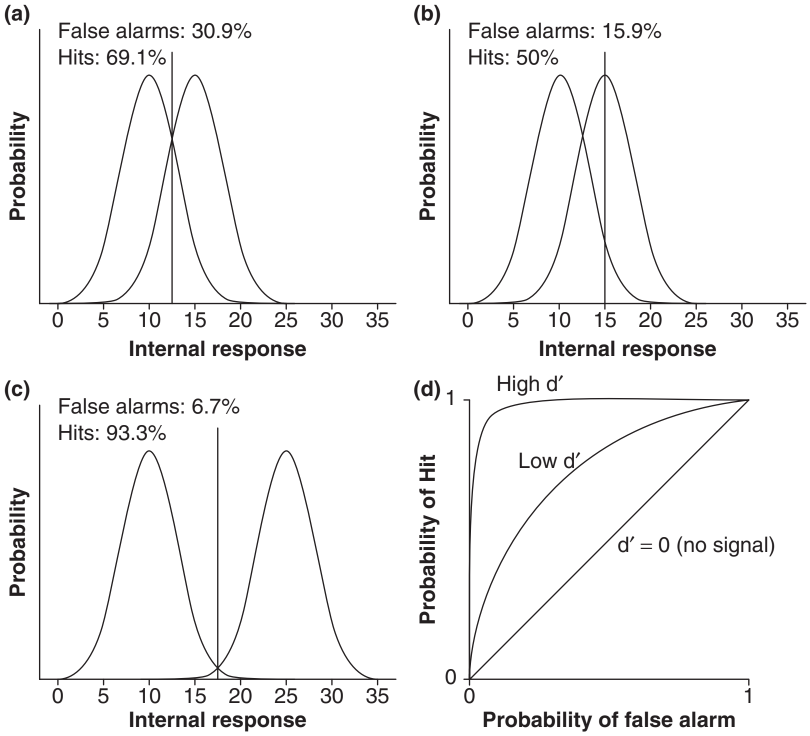


Figure 2 Discriminability and the ROC Curve

Notes: (a) Internal response and possible outcomes for the case of a symmetric criterion. (b) A higher criterion decreases the likelihood of both hits and false alarms. (c) As signal strength is increased, decision makers can reduce both false alarms and misses. (d) Receiver operating characteristic (ROC) curves for three different signal strengths (indexed by discriminability or d′) showing the combinations of hit and false alarm probabilities consistent with a given level of perceptual discriminability.

The outcome depends critically on the placement of the criterion. A symmetric criterion (Figure 2a) results in a fairly high probability of a hit (large area to the right of the criterion), but a substantial false alarm probability as well. Increasing the criterion (Figure 2b) corresponds to requiring better evidence (greater internal response) before saying "yes" (i.e., it is a more conservative criterion). The result is that a false alarm is less likely, but at the cost of reducing the likelihood of a hit as well. Thus, the position of the criterion corresponds to the decision maker's bias to say "yes" or "no."

For this case, there is no way for the decision maker to set the criterion to achieve only hits and no false alarms. The noise can never be avoided. The internal response for the signal plus noise case is generally greater but there is a distribution (a spread) of possible internal responses. Because the two curves overlap, the internal response for noise alone might exceed the internal response for signal plus noise. Because of this, the decision maker cannot always be right. The decision maker can adjust the kind of errors that he or she makes by manipulating his or her criterion, the one part of this diagram that is under the decision maker's control.

Increasing signal strength by providing more information (e.g., a better quality CT or MRI scan) makes the decision easier. This corresponds to a greater separation of the two curves (Figure 2c), allowing for an increase in the likelihood of a hit, while simultaneously reducing the likelihood of a false alarm (compare Figures 2a and 2c).

The Receiver Operating Characteristic

The full range of a decision maker's behavioral performance, for all possible settings of the criterion, can be captured in a graph, called the *receiver operating characteristic* (ROC). ROC curves (Figure 2d) are plotted with the probability of a false alarm on the horizontal axis and the probability of a hit on the vertical axis. The figure shows three examples of ROC curves, corresponding to the two signal strengths in Figures 2(a) and 2(c), as well as a signal strength of zero. Just pay attention to the middle (Low d′) ROC curve for the time being. If the criterion is very high, then both the false alarm and the hit probabilities will be very low, near the lower left corner of the ROC graph. If the criterion is very low, then both the hit and the false alarm probabilities will be very high, near the upper right corner of the graph. For an intermediate choice of criterion, the hit and false alarm probabilities will take on intermediate values. The decision maker may set the criterion anywhere, but any choice will land them with a hit and false alarm probability somewhere on the ROC curve for that particular signal strength. Notice also that for any reasonable choice of criterion and nonzero signal strength, the hit probability is always larger than the false alarm probability, so the ROC curve is bowed upward.

With more information, there is more separation between the two probability-of-occurrence curves, and the decision maker can pick a criterion to get a very high probability of a hit with almost no likelihood of a false alarm (Figure 2c). ROC curves for stronger signals bow out further toward the upper-left-corner of the graph than ROC curves for weaker signals (Figure 2d, High d′). Ultimately, if the signal is really strong (lots of information), then the ROC curve goes all the way up to the upper left corner (perfect performance with a 100% chance of a hit and no possibility of a false alarm).

The shape of an ROC curve can be summarized by a single number called the *discriminability index* or d′ (pronounced “dee prime”). The discriminability index captures the inherent difficulty of the decision maker’s task, independent of his or her criterion. A hard task with only a weak signal strength yields an ROC curve near the diagonal line (Figure 2c, d′ = 0) and an easy task with a strong signal yields an ROC curve that bends up to the upper left corner (Figure 2c, High d′). The primary virtue of d′ is that its value does not depend upon the decision maker’s criterion, but instead it is a true measure of the information content in the internal response.

The decision maker’s criterion will generally depend on a cost-benefit analysis of the task. For example, decision makers will typically adopt a more liberal (lower) criterion if the *payoffs* are altered to benefit “yes” answers (higher payoffs for a hit compared to a correct reject, lower penalties for a false alarm compared to a miss). This same lowering of the criterion can also happen in response to a change in the *prior probabilities* (if signal present is known to occur more frequently than signal absent). Finally, if the signal strength is increased (shifting the signal plus noise curve to the right as in Figure 2c), the decision maker will typically shift the criterion to the right, increasing the probability of a hit and decreasing the probability of a false alarm. But it is difficult to choose an optimal criterion value, and people sometimes struggle with it. When targets are extremely rare (for example, searching baggage at an airport), criteria are typically set high, resulting in a large chance of misses, despite the far higher cost of a miss as compared to that of a false alarm.

Measuring Discriminability and Criterion

In a perception experiment, signal strength is typically under the control of the experimenter, and the likelihood of a hit or false alarm is measured for each signal strength. For example, in a hearing test, a person’s threshold for hearing is measured by adjusting the amplitude of a sound until it is just barely detectable. In the laboratory, the forced-choice protocol is the preferred experimental procedure. A sound is played on half of the trials, but no sound is played on the other half of the trials. The subject is forced to respond on every trial either “Yes, I heard a sound” or “No, I didn’t.” Because of internal noise, the subject sometimes presses the “no” and sometimes the “yes” button for the same stimulus on subsequent trials. Typically, several different signal strengths (loudnesses) are interleaved. For each signal strength, the hit rate is the percentage of signal-present trials in which the subject responded “yes.” The false alarm rate is the percentage of signal-absent trials in which the subject responded “yes.”

Discriminability (d′) and criterion are then estimated from the hit and false alarm rates.

With both the hit rate and the false alarm rate in hand, d′ is simply determined by noting on which of a family of ROC curves the performance lies (Figure 2d). Criterion can also be determined by where along the ROC curve the performance falls.

David J. Heeger
and Michael S. Landy

See also Audition: Loudness; Auditory Thresholds

Further Readings

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Huntington, NY: Krieger.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.

Wickens, T. D. (2002). *Elementary signal detection theory*. New York: Oxford University Press.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual search. *Nature, 435*, 439–440.

SIZE PERCEPTION

See Constancy

SKIN SENSES

See Cutaneous Perception

SLEEP AND DREAMS

The visual system is highly dependent on sleep for the organization of the brain and learning about the visual world. During the first year of life, one of the steepest periods of experience-dependent learning, infants sleep 70 to 80% of the time each day. Sleep deprivation during critical periods of kitten visual development produces abnormal organization of the visual cortex. These studies indicate that normal development of perception and the brain structures subserving perception are dependent on early experience and sleep. Although sleep needs changes throughout the lifespan, research shows that sleep is essential for health, memory, and restorative processes, including visual learning.

Dreams during rapid eye movement (REM) sleep are considered perceptual processes without the constraints of external stimuli. Dreams are multisensory experiences for both sighted and congenitally blind individuals. This indicates that the ability to form visual images may be independent of visual perception. This entry focuses on sleep basics, dreams and experience, visual learning and nocturnal sleep, and visual learning and naps.

Sleep Basics

To discuss the role of sleep for perception, it is helpful to introduce some basic concepts. Sleep is a highly structured set of processes separated into five stages, each demonstrating: (a) stereotypic electrical activity, (b) neurochemical expressions, and (c) both enhancement and depression in varying brain regions. The five stages (stages 1, 2, 3, 4, and REM) progress in a cycle from stage 1 through stage 4 and then to REM sleep. The duration of an entire cycle lasts for 90 to 110 minutes. Adults spend 60% of sleep in stage 2, about 20% in REM, and the remaining 20% in stages 3 and 4, which comprise slow-wave sleep (SWS). Infants spend about 50% of sleep in REM—an observation cited as evidence for the importance of REM in the developing brain. Stage 2 sleep is characterized by fast 12 to 14 hertz (Hz) waves (called spindles) and slower K complex waves. SWS consists of extremely slow brain waves, called delta waves, interspersed with smaller, faster waves. REM sleep, in contrast to SWS, is a lighter sleep accompanied by rapid irregular shallow breathing, rapid, jerking eye movements, increased heart rate, increased cortical blood flow, limb muscle paralysis, and a predominance of theta waves.

Sleep cycles vary systematically during the night. Specifically, the first part of the night is dominated by SWS. As the night progresses, a reciprocal increase in REM sleep displaces SWS. The morning period is, thus, rich in REM sleep. In research, sleep is commonly divided into REM and NREM (non-REM) sleep.

Dreams and Experience

Most dreams occur during REM sleep. Stephen LaBerge and Howard Rheingold have written that dreams can be viewed as a special case of perception without the constraints of extrasensory input. Conversely, perception can be viewed as the special case of dreaming constrained by sensory input. During REM, sensory input from the external world and body movement are suppressed, but the brain, especially the areas associated with perception, is highly active. Neural activity in the sensory areas allows the dreamer to see (in color), hear, feel, and even taste things that are not present in the external environment.

Dreams are useful for studying perceptual phenomena, such as visual imagery (i.e., the ability to hold an internal visual representation of a thing that is not present in the external world). An interesting question is whether visual imagery depends on visual perception, or whether imagery emerges from activation of the visual cortex by nonvisual inputs. One way to examine this question is by investigating the degree of visual content in the dreams of individuals who have not had any visual experiences (e.g., congenitally blind individuals). Helder Bértolo and colleagues analyzed EEG alpha power in congenitally blind subjects, as a possible indicator of the visual content of their dreams, and evaluated the ability of congenitally blind to graphically represent the dream-evoked images. He found strong evidence that congenitally blind individuals are not only able to verbally describe what may be the visual content of their dreams, but they can provide, through drawing, a graphical representation of such content. Also in the study, a significant negative correlation between the visual content of the dreams and the alpha power was found in both blind and nonblind groups. These data call into question the assumption that experience is essential both for visual imagery and for visualization. Visualization without previous experience, as is the case for congenitally blind individuals, would indicate the existence of visual imagery independent of visual perception. This implies that the born-blind subjects are capable of using other sensory modalities to integrate these inputs via the visual system to produce concepts capable of graphical representation.

The most compelling evidence of a relationship between waking experience and dreams comes from studies of sleep onset dreaming, called hypnagogic dreams. During sleep onset, hallucinations increase, whereas directed thinking decreases. Importantly, these truncated dreams show robust incorporation of daytime experiences. Robert Stickgold and colleagues manipulated hypnagogic dream content by having subjects play the video game Tetris or the arcade-style downhill skiing simulator, Alpine Racer. As subjects fell asleep, reports were collected about their dream content. Using these games, the researchers obtained sleep onset reports of images of Tetris or downhill skiing in up to 89% of subjects and 42% of first night reports. Surprisingly, they found no difference in frequency or content between normal and densely amnesic subjects. Nevertheless, the neocortical sources of these images were not simply stored sensory representations of recent stimuli, as Tetris players occasionally reported images from past versions of Tetris and Alpine Racers reported images from actual skiing. Although the sleep onset period differs from normal NREM and REM sleep, these findings, nevertheless, make an exciting contribution to models explaining the nature and function of dreaming by showing that dreams are an amalgam of recent and older events that become incorporated into our cortical associative networks.

Visual Learning and Nocturnal Sleep

REM and SWS contribute differently to memory; REM being more important for *procedural memory* (e.g., knowing "how," learning actions, habits, perceptual and motor skills, and implicit learning), whereas SWS appears to more strongly support *declarative memory* processing (e.g., knowing "what," in terms of events, places, and general knowledge). Some of the clearest evidence of sleep-dependent learning comes from visual perceptual learning studies that utilize the texture discrimination task, in which subjects identify the orientation of an array of lines that flash very briefly in the periphery. Perceptual learning is the specific and relatively permanent modification of perception and behavior following sensory experience. Avi Karni and colleagues have shown that post-training improvement on the texture discrimination task is only evident several hours after training, and that improvement can develop overnight, although only when REM sleep is allowed. Extending these findings, researchers demonstrated the following: (a) that improvement in performance on the texture discrimination task can be achieved only after six hours of nocturnal sleep; (b) additional nights of sleep appear to produce additional, incremental improvements in performance. These improvements (a) occurred even in the absence of additional training; and (b) demonstrated sleep-stage dependency.

Performance improvements correlated with the amount of SWS and REM individual subjects had during nocturnal sleep. No significant correlation was seen between SWS and REM, indicating that each contributed independently to the improved performance. Along with improved performance, sleep-dependent perceptual learning leads to enlarged regions of activation in the primary and secondary visual cortex measured with functional magnetic resonance imaging. Further, electrophysiological studies showed visual field specific, evoked potential enhancements in the occipital C1 component that corresponded to the trained visual quadrant.

Visual Learning and Naps

Sara Mednick and colleagues have shown that daytime naps rich in both SWS and REM produce perceptual learning and ameliorate the detrimental effects of visual fatigue due to overpractice. The researchers showed that same-day, repeated testing on a visual task decreased performance in a visual field–specific manner. Interestingly, naps with only SWS prevented this perceptual deterioration and maintained performance at baseline levels across the day. Additionally, naps with both SWS and REM actually showed significant improvement from baseline. This nap-dependent improvement was similar to improvement following a night of sleep in its magnitude, visual field specificity, and sleep stage dependence. Similar to nocturnal sleep studies, nap-dependent learning was additive with a following night of nocturnal sleep, such that morning performance after a nap plus nocturnal sleep was significantly better than after only nocturnal sleep and equivalent to the performance of subjects who retested after two nights of sleep.

Sara C. Mednick

See also Perceptual Learning; Texture Perception: Tactile; Video Games; Visual Disorders: Blindness

Further Readings

Bértolo, H., Paiva, T., Pessoa, L., Mestre, T., Marques, R., & Santos, S. (2003). Visual dream content, graphical representation and EEG alpha activity in congenitally blind subjects. *Cognitive Brain Research, 15,* 277–284.

Dadvand, L., Stryker, M. P., & Frank, M. G. (2006). Sleep does not enhance the recovery of deprived eye responses in developing visual cortex. *Neuroscience, 143,* 815–826.

Frank, M. G., Issa, N. P., & Stryker, M. P. (2001). Sleep enhances plasticity in the developing visual cortex. *Neuron, 30,* 275–287.

Mednick, S., Nakayama, K., & Stickgold, R. (2003). Sleep-dependent learning: A nap is as good as a night. *Nature Neuroscience, 6,* 697–698.

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature, 437,* 1272–1278.

SOCIAL PERCEPTION

Social perception refers to the broad tendency of perceivers to note and interpret the appearance, behavior, and intentions of others. Such perception is informed by multiple cues that originate in body features and behaviors, and it affects the perceivers' attitudes about and behaviors directed toward others.

Historically, the phrase social perception meant very different things, depending on a researcher's background. Within the vision and cognitive sciences, for example, social perception research focused heavily on how observers process cues that are diagnostic of social factors (e.g., face perception and biological motion perception). Within social psychology, in contrast, social perception research focused primarily on the implications of perceiving meaningful social categories (e.g., stereotyping and prejudice). Thus, depending on one's research perspective, social perception could mean very different things. In spite of these isolated beginnings, social perception research is experiencing a surge of unprecedented interdisciplinary focus. This entry explores the topics of social perception within vision and cognitive sciences, social perception within social psychology, and social vision as an emerging hybrid research model.

Social Perception Within Vision and Cognitive Sciences

Arguably the most meaningful stimuli in our environment are other people. We see others with great regularity, and our perceptual systems appear to be remarkably well suited to interpreting the extant visual cues. Understanding the mechanisms that underlie the perception of such cues has been a longstanding question for vision and cognitive scientists, and research has focused heavily on perception of the face and body.

Face Perception

The face garners considerable attention from others. The face is likely to be the first place observers look when they encounter another person. And the propensity to examine the face is evident early in life. At birth infants show marked preferences to look toward facelike stimuli. Given the overwhelming tendencies for people to look toward the face, perhaps it is unsurprising that extensive research has been devoted to understanding face perception.

What type of information is furnished by the face? Perhaps the single most important thing that face perception serves is identity recognition. Perceiving the particular spatial configuration of facial features enables observers to determine whether a face belongs to one's best friend or their eccentric professor. Other entries in this volume describe in greater detail the cognitive and neural mechanisms that govern face perception (see the list at the end of this entry).

Another class of information that is readily extracted during face perception is emotional state. Emotion categories correspond to distinct patterns of contraction in the facial muscles. Beginning with Paul Ekman's pioneering work, these patterns are now well documented. Much of this work suggests that emotion expressions are not only easily recognized by observers, but that the expression and perception of six basic emotions is universal across cultures. These emotions include sadness, happiness, anger, disgust, fear, and surprise.

A final class of information that is appreciated during face perception is one's social category membership. The internal structure of the face provides reliable cues to social categories such as sex, race, and age, and observers infer social categories from a person's facial cues. The interesting ways in which these factors interact with one another are discussed further in this entry.

Body Perception

The face is not alone in its ability to provide observers with meaningful social information. Indeed, it is frequently the case that social perception occurs at a physical distance or a visual vantage that precludes face perception. Under such circumstances, social perception proceeds unencumbered, relying on body cues as a source of information.

The body provides two types of information. The first type of information provided by the body is its shape, and this conveys many things to observers. Because men and women vary in both absolute and relative measurements, such cues are reliable indicators of sex category membership. Additionally, the body's posture can convey other social information, such as emotion state.

The second source of information provided by the body is its motion. The body's motion is sufficient to inform a variety of judgments, including the perception of personal identities, social categories, and emotion states. From such information, observers may form the impression regarding another person's attractiveness, hostility, or vulnerability to attack. Under some circumstances, the perception of body shape may constrain the interpretation of body motion.

Though the focus in this entry so far has been on visual perception, vision is not the only means by which social perception occurs. Much of the information that is perceived visually can also be discerned acoustically. Qualities of the voice, for example, provide reliable cues for identity recognition and social categorization (including sex, race, age, and even sexual orientation). Similarly, the sound of footsteps can be sufficient to inform judgments of categories and identities.

To summarize, within the vision and cognitive sciences, social perception research has sought to understand how a variety of embodied cues are perceived by observers.

Social Perception Within Social Psychology

Attributional Approaches

Within social psychology, the phrase *social perception* has long been synonymous with *person perception,* used to describe the tendency of observers to discern others' traits and dispositions. Early researchers in this area sought to understand whether stable individual differences in personality could be judged by observers. Because several of the factors that are most important to social situations (such as loyalty, honesty, and agreeableness) are not directly observable, the perception of such factors had to be inferred from behavior. This fact led to uncertainty about the validity of observers' judgments because accuracy could not be verified objectively. Though early researchers were frustrated by this fact, other work began to focus on a more manageable question—the conditions under which observers of behaviors will draw inferences about a person's disposition, rather than attributing actions to situational determinants.

Fritz Heider was among the first to recognize that an observer faces a dilemma of whether to attribute a person's actions to internal factors (e.g., personality, disposition) or external circumstances (situational constraints). If an observer witnessed Professor Smith dismissing an upset student in the hallway, for example, the behavior could be attributed to an internal factor (that Professor Smith is uncaring) or to an external factor (Professor Smith was rushing out of the classroom because he was late). Such basic observations provided the foundation for two related theories. Harold Kelly's *covariation model,* for example, described how people judge the attitudes of others based on their behaviors. Edward E. Jones and Keith Davis's *theory of correspondent inferences* described why observers are prone to attribute behaviors to dispositions. Across dozens of studies, researchers documented that people are prone to draw dispositional inferences after viewing the actions of others.

Although person perception research began with a quest to understand the type of attributions that people make after observing the behavior of others, contemporary social perception research is more varied. Many of the social categorization judgments that vision and cognitive scholars studied to understand how cues are perceived are a foundation for interpersonal phenomena. Indeed, a large body of work has examined both the inevitability and consequences of social categorization.

Perceiving Social Categories

Since Gordon Allport's early work, social categorization has been theorized to be an efficient cognitive strategy that streamlines interpersonal interactions. It was argued that knowing another person's social category membership enabled observers to recruit knowledge structures that related to the perceived category (i.e., stereotypes). Once activated, stereotypes can have a profound impact on interpersonal attitudes and behaviors.

Although a person may fall into numerous social categories, only a small subset of categories is highly likely to be perceived by observers. These include sex, race, and age. These categories are readily perceived by visual cues that are apparent in the faces and bodies of others. Their perception requires so little effort that some scholars have argued that the process is obligatory and automatic.

Regardless of whether social categorization is supported by face or body perception, it carries significant interpersonal consequences. Sociologists, for example, recognized the impact of social categorization, describing categories such as sex, race, and age as *master status categories.* This term implies that once known, some categories have an irrevocable impact on subsequent interactions.

Consequences of Social Perception

Social category knowledge caries interpersonal consequences because it evokes related knowledge structures, or stereotypes. Categorizing another person according to their sex, for example, also brings to mind sex-based stereotypes of men or women (e.g., that men are competitive or that women are nurturing). Although this occurs outside of awareness, the activation of these stereotypes affects the perception of unrelated factors, such as the evaluation of competence in the workplace or memory for an individual's appearance or actions.

The perception of others and the activation of related stereotypes also lead perceivers to change their own behaviors. The stereotypes of many social categories contain behavioral components. Examples include: Men are competitive; Women are emotional; Elderly people walk slowly. Stereotypes such as these are widely known, even though they are not universally believed. Even if one does not endorse the stereotype generally, or the specific element of the stereotype specifically, it can affect behavior. Numerous studies have demonstrated such effects. For example, activating the elderly stereotype leads people to walk more slowly; activating the black stereotype leads people to behave more aggressively; and activating the professor stereotype leads people to perform better on standardized tests. Thus, perceiving social categories has implications not only for impressions of the perceived individual, but also for the behaviors of the perceiver!

Social Vision: An Emerging Hybrid Research Model

The previously described research reviewed social perception research from two distinct theoretical and empirical traditions—vision and cognitive sciences on one hand and social psychology on the other. The historic treatment of social perception from each of these perspectives covers decidedly different terrain in the social perception landscape. Vision and cognitive scholars were most interested in the perception of particular cues that resulted in either accurate or inaccurate social judgments. Social psychologists, in contrast, investigated social perception with categorization as a given, and explored the downstream consequences thereof. Thus, answering the question, "What is social perception" was far from straightforward. It required different answers for different disciplines, as previously reviewed. Recently, however, social perception research is becoming increasingly interdisciplinary, and such research has been described as *social vision.*

One area of social vision research has examined the perception of "thin slices" of behavior, or judgments based on very little information. For decades, vision scholars have been able to quantify the efficiency of social perception because the social categories of interest were discrete (e.g., man or woman, young or old). Researchers in social psychology, in contrast, puzzled over how to objectively quantify the accuracy of social perception for domains that were not objectively defined (e.g., personality traits). New research has overcome this obstacle by providing research participants with minimal visual information, or thin slices of behavior, to judge on dimensions that are either continuous or difficult to discern. Such studies have included judgments of effectiveness for silent degraded videos (30 seconds) of professors lecturing to a class of students, judgments of sexual orientation for silent video images depicting a person talking (2 seconds) or walking on a treadmill (10 seconds), judgments of leadership ability for photographs of unknown political candidates, and judgments of power for photographs of Fortune 500 CEOs. When compared to more objective indices, such judgments show a surprising degree of accuracy. Research participants' judgments closely mirrored objective measures: end of term course evaluations, self-identified sexual orientation, actual election outcomes, and corporate profits, respectively.

In much of the thin slices research, the specific cues that convey information are not entirely clear. In other research, however, care has been taken to quantify the parameters of face and body cues, and these are related to the accuracy of perceptions. The adage "you can't judge a book by its cover," when it comes to some domains, appears questionable, at best.

Another example of social vision research examines the social constraints on visual perception.

This happens in at least two ways. First, the perception of one cue to social perception is constrained by another cue. For example, if an observer already knows the sex category of an individual, it is likely that sex-specific cues in the face or body will be perceived and evaluated within the context of the category. If the cues are consistent, evaluative judgments (such as attractiveness) will be favorable. If the cues are inconsistent, the evaluation will be unfavorable. Second, knowledge of existing stereotypes can constrain the perception and interpretation of visual cues. The perception of emotional body motion, for example, can bias the perception of sex category. Displays of emotion are highly sex stereotyped—women are presumed to react to challenges with sadness, whereas men react with anger. When participants judge the sex category of display emotional body motion merely as points of light cast against a dark background, their judgments reflect such stereotype knowledge. When these displays depict anger, they are overwhelmingly judged to be men; when these displays depict sadness, they are judged to be women. Put simply, social factors constrain the perception of visual cues.

Thus, traditional research within social psychology began where the vision and cognitive scholars' work ended. Increasingly, however, it has become apparent that social cognitive processes can modulate the perception of social cues.

Kerri Johnson

See also Face Perception; Motion Perception: Social

Further Readings

Adams, R. B., Jr., Ambdy, N., Nakayama, K., & Shimojo, S. (2008). *The science of social vision.* New York: Oxford University Press.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111,* 256–274.

Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 89–150). New York: McGraw-Hill.

Gladwell, M. (2005). *Blink: The power of thinking without thinking.* New York: Little, Brown.

Knoblich, G., Thornton, I., Grosjean, M., & Shiffrar, M. (2005). *Human body perception from the inside out.* Cambridge, MA: Oxford University Press.

Peterson, M., & Rhodes, G. (2003). Advances in visual cognition. *Perception of faces, objects, and scenes: Analytic and holistic processes* (Vol. 2). Cambridge, MA: Oxford University Press.

SOMATIC PERCEPTION

See Body Perception

SOUND REPRODUCTION AND PERCEPTION

Imagine you've gone to hear your favorite musician in concert. After the concert is over, you'd like a souvenir (from the French word for "remember") of the evening. Today, that auditory souvenir would be a recording, most likely a CD or MP3 of the performance. You would want it to sound as much like the live show as possible, to capture the emotional qualities of the performance as well as the sound of the concert hall, the spatial image of the sound, and the unique sound of the instruments—collectively, all aspects of the fidelity of the recording (from the French *fidélité:* accuracy and faithfulness). Not all sound recording has fidelity as its goal, however. In recent decades, artists and engineers have experimented with sounds that could never be produced in a natural setting, pushing the boundaries of artistic expression. This entry surveys the various ways of reproducing sound, from the early piano rolls to digital technologies; considers the role of delivery systems such as loudspeakers and headphones on the listening experience; and describes the debate regarding the superiority of analog versus digital reproduction.

Modern attempts to preserve musical performances began in the mid-1800s with the invention of piano rolls. These long rolls of paper contain precisely spaced holes that are read by a mechanism causing the hammers to hit the piano strings as though a pianist were pressing the key. Before the piano roll, if you wanted to hear music, you either had to play it yourself or find a musician who could, and every musical performance was subtly different, subject to the limitations of human motor control, if not to differences in the performer's emotional state. The piano roll introduced two innovations: *music-on-demand,* and the notion of a *master* or ideal performance. (These were subsequently refined and instantiated on various technologies, such as wax cylinders, magnetic tape, acetate, vinyl, and digital sound files.) Music-on-demand meant for the first time listeners could hear music whenever they wanted to, without a performer physically present, and could hear performances from musicians who were in another country or even dead. Moreover, knowing that a defect might be preserved for posterity, and listened to over and over again, musicians were driven to create better performances than they might have otherwise, while technicians developed ways to correct imperfections in the rolls themselves in order to improve the playback.

Transducing horns were subsequently used to capture the acoustic waveforms of a performance on wax cylinders. For singers in particular, this required projecting loudly enough to be heard over the band, or screaming right into the horn itself. With the development of microphone technology in the 1930s and 1940s, singers could sing directly into an efficient transducer, allowing for the refinement of vocal performances. For the first time, a vocalist could sing softly—even whisper—and be heard over even a loud band. This difference is apparent in the premicrophone recordings of Bessie Smith versus the miked recordings of Billie Holiday.

Piano rolls represented an early attempt to create an artifact that contained fewer flaws than the actual performance that it was based on. Microphones created a recording that was neither an accurate reproduction of what occurred during the performance nor what could actually be performed—that is, they were among attempts to create an artificial soundscape for aesthetic purposes. Close-miking techniques, in which a microphone is placed right next to individual instruments rather than over the musical group as a whole, create such *hyperrealities,* allowing a listener to imagine that his or her ears are right in front of each musical instrument simultaneously, to hear acoustic nuances that normally would be masked by other instruments. Such techniques reached their peak in the 1960s and 1970s recordings by Stevie Wonder, the Beatles, and Steely Dan (among others), in which each instrument can be heard more clearly and distinctly than could ever occur in a live setting.

Stereo recording introduced three-dimensional acoustic imaging. Modern mixing techniques allow for the positioning of a musical instrument uniquely in the left-right plane as well as a front-back plane (using reverberation), adding even more individual distinctiveness to each instrument. By the 1960s, recording engineers were increasingly employing analog signal processing devices (and by the 1990s digital signal processing, or DSP, was dominant). Signal processing permitted more careful control of the recorded sound. It was first used to compensate for recording artifacts, such as phase and spectral imperfections that subtly change amplitude and timing relationships between the different frequency components of a complex sound, and later as effects in themselves. These include compression, expansion, equalization, gating, and artificial reverberation and delays. Compression, expansion, and gating allow control of amplitude relationships, treating loud and soft sounds differently, and may be used to make vocals more intelligible or instruments more prominent in a mix. Equalization allows the manipulation of the frequency (spectral) content of a sound; the bass and treble controls on the stereo are an example. Additional special effects such as flanging (combining a variably delayed copy of a signal with itself, producing audible spectral changes) and chorusing (artificial duplication using multiple varying delays and amplitude variations) were made popular by the Beatles.

The dominant form of professional recording between the 1930s and 1980s relied on analog, magnetic tape, in which amplified electrical signals from microphones were converted to magnetic field variations that could be stored as magnetic patterns. Magnetic tape consists of many tiny magnetic particles glued to a strip of plastic that is moved past a recording head at constant speed. The head generates a magnetic field *analogous* to the electrical signal from the microphone (hence the term *analog recording*) strong enough to magnetize the tape particles so they retain the pattern for later playback. Digital recorders (commercially available beginning in the late 1970s) convert the continuous analog signal to a list of numbers sampled at a high rate and were hailed by many audiophiles as providing cleaner sound because they eliminated the tape hiss and speed variations associated with analog tape. The introduction of the digital compact disc in 1983 allowed for the digital encoding of any recording, whether it began as an analog or a digital signal. It is now widely believed that the particular sonic parameters chosen for the commercial CD entailed a compromise in quality, such that the sampling rate and bit rate are not capable of yielding truly high fidelity; indeed, many vinyl records have a higher frequency response.

The type of transducer used to reproduce the music plays a role in the listening experience and influences how music is mixed and produced as well. Loudspeakers interact with their acoustical environment so that each room and system differs from every other, leading to what has been called *the second venue problem:* the idea that the original room or concert hall in which a recording was made must necessarily interact with the acoustics of the room in which the recording is being played back. Headphones and earbuds remove the influence of the room entirely and deliver subtle details, such as reverberation decays, which are sometimes lost on loudspeaker systems, but the sounds appear to be localized inside our heads rather than in front of us in space. Special "dummy head" microphones can be used to capture realistic spatial imaging for headphone playback, but the technique fails when played on loudspeakers. In general, the halfway-in (nonoccluding) ear buds tend to lack the low frequency response that is provided by in-the-ear inserts or over-the-ear headphones.

An active debate among audiophiles concerns the inherent superiority of analog versus digital recording and reproduction. A common misconception is that CDs provide sound that is "purely digital," but loudspeakers and headphones are analog devices (they are constructed from components that physically move the air molecules in order to create sound waves) and so the end stage of digital reproduction always depends on digital to analog converters (DACs), and their quality can vary widely. While a scientific analysis of these competing systems gives digital the advantage, listeners' preferences do not always reflect this, possibly due to imperfect "real-world" implementations. The difference between digital and analog sound is much like the difference between digital video and film images: the digital media capture sharp edges and high contrast, and the analog media gently blur edges and soften images.

"True" or "controlled" experiments are required in order to evaluate the many claims of equipment manufacturers and audiophiles. Holding extraneous variables constant is critical to such experimental designs. Ideally, a listener should evaluate two sets of loudspeakers using identical amplifiers, wiring, and source material, and statistical analysis (such as the binomial or Chi-Square test) will determine if the results are better than random guessing.

Many claims of equipment manufacturers are unproven, and many more lack any scientific basis. Examples of the latter include claims that a high quality power cord will enable a CD player to spin at a more accurate speed and thus yield higher audio quality (this is not necessary because servo units accept a wide range of voltages and currents while maintaining constant speed, and moreover, speed fluctuations do not have audible consequences when reading digital information), or that placing magnets or pyramids on top of components will "channel" or "focus" the current better.

Daniel J. Levitin and Jay Kadis

See also Acoustics and Concert Halls; Auditory Scene Analysis

Further Readings

Levitin, D. J. (1999). Experimental design in psychoacoustic research. In P. R. Cook (Ed.), *Music, cognition and computerized sound: An introduction to psychoacoustics* (299–328). Cambridge: M.I.T. Press.

Rumsey, F., & McCormick, T. (2006). *Sound and recording* (5th ed.). Oxford, UK: Focal Press.

SOUND STIMULUS

Our entire scientific knowledge of how hearing works derives from the stimuli used in auditory experiments. Hearing scientists have devoted much effort and insight into designing the sounds for their experiments so that each is just right for its purpose. This has become much easier in recent years, as the near-universal use of computers means that almost any stimulus that can be thought of can be used, provided only that a computer program can be written to make it. The result is a true revolution in experimental method: The days when specialized electrical hardware was needed for auditory experiments—and which only a few laboratories had—are long past.

The three classic stimuli of auditory science discussed in this entry are pure tones, noises, and clicks. Their popularity comes from the ease with which they can be generated, be it electrically or computationally, but also from their simple properties: They are the building blocks underlying all other stimuli. Indeed, it is rare to encounter an experiment that does not use at least one of them.

The pure tone is the simplest. Its only physical parameters are frequency, measured in Hertz; power, measured in decibels; and duration, measured in seconds. The first two parameters are closely related to the perceptions of pitch and loudness: A pure tone has fixed, unchanging pitch and loudness (indeed, they are often taken as references when measuring the perceptions of more-complex stimuli). Pitch is particularly important: A pure tone has only a single frequency component and evokes only a single pitch, but combinations of pure tones may have one or more pitches. The search for the rules underlying which particular pitch results from a pure tone or a set of pure tones has driven much of the work on theories of pitch perception. A "white" noise, in contrast, contains every frequency: It is often made by adding together every possible pure tone, each with a random level and phase (the name "white" is by analogy with white light, which contains all visible frequencies). Its spectrum is broad and evokes no pitch. Countless experiments have used white noises as masking sounds that measure the detection threshold of some other sound in the noise. Countless others have shaped the spectrum of the noise in some way to give maskers with special characteristics, such as by giving proportionally less weighting to higher frequencies than lower (e.g., "pink" noise) or by filtering it to include only certain frequencies (e.g., "notched" noise).

When used in experiments, tones and noises can last from a few milliseconds (ms) to a few minutes. In contrast, the shortest stimulus that can be made is a click, usually with a duration less than 0.1 ms. But a single click is rarely used in experiments, as it is just too short. Instead, experiments usually use sets of them, from two-click pairs to multiclick trains lasting a second or more.

The definite pitch and rich timbre produced by musical instruments are due to the large number of pure tones they generate. The frequencies of these "harmonics" are often multiples of the lowest harmonic (i.e., if the fundamental frequency measured in hertz [Hz] is at f Hz, then the rest are at $2f$, $3f$, $4f$, $5f$, and so on), but the intensity of each one will vary across instruments. This pattern helps to determine the timbre of the sound and distinguishes one instrument from another (e.g., oboes give very little power at the fundamental frequency f, while clarinets give very little power at the even-numbered harmonics $2f$, $4f$, etc.).

The dynamic changes of the amplitude of the sounds generated by different musical instruments are also characteristic for determining their perception. In fact, most hearing takes place with sounds that vary: most real-world sounds have amplitudes and frequencies that vary with time (for example, the wail of a police siren is due to a repeated sweep in frequency). Natural speech is an extremely modulated stimulus, and much of the prosodic and phonetic information is carried by slow or fast fluctuations in the power at different frequencies. Modulations are also important for hearing aids, as they often react to varying sounds and change their behavior accordingly, for instance by decreasing their amplification if the input sound suddenly gets more intense.

With computer-generated stimuli, a stimulus can be made with modulations that are predictable or random, shallow or deep, smooth or jagged, and which may or may not be the same at different frequencies. It is also possible to extract the modulations of one sound (e.g., speech) and apply them to another, completely different sound (e.g., music). These *auditory chimaeras* have a mixture of perceptual properties, some from the speech and some from the music. Determining which is which can determine much about how the auditory system works. An analogous approach—removing as many of the modulations as possible from speech—is a new technique and showing much scientific promise. All these complex stimuli are only achievable through special computer programming and exemplify the state of the art in stimulus generation.

Michael A. Akeroyd

See also Audition; Audition: Pitch Perception; Audition: Temporal Factors; Hearing Aids; Psychophysical Approach; Sound Reproduction and Perception; Speech Perception

Further Readings

Hartmann, W. M. (1997). *Signals, sound, and sensation.* New York: American Institute of Physics/Springer-Verlag.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2005). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences USA, 103,* 18866–18869.

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception, *Nature, 416,* 87–90.

Yost, W. A. (2007). *Fundamentals of hearing: An introduction.* San Diego, CA: Academic Press.

SPATIAL LAYOUT PERCEPTION, NEURAL

The ability to perceive the layout of objects around us is so critical for survival that individuals who are unable to do so are virtually helpless. Understanding the neural underpinnings of spatial layout perception aids in predicting and dealing with the consequences of brain injuries and also provides an important framework for understanding how the brain builds representations of object locations under more normal circumstances. A majority of the brain is devoted to perceiving and making use of spatial information in some way. In the case of vision, visual information flows from the eyes, through the thalamus, and into the primary visual cortex; conscious perception of spatial layout is thought to be processed in neural pathways running from the visual cortex to the inferotemporal cortex and into the medial temporal lobe. Relatively little is known about the neural basis of specific forms of visual layout information. One important exception is that binocular (or retinal) disparity, a strong depth cue involving signals from the two eyes, is processed in the primary visual cortex—relatively early in visual processing. Motion-based depth cues, meanwhile, are processed in the V5/MT complex. This entry discusses what constitutes perception and the foundational processes of perception, as well as describing the neural basis of spatial perception in vision and audition.

What Is "Perception"?

There are many aspects of our experience of spatial layout. For example, on a clear, dry day, a distant mountain may look relatively close—perhaps only 10 kilometers (km) away. However, our past experience (or a map) may tell us that the mountain is really 50 km away. If we are stuck in traffic and in a hurry to get there, the mountain may seem to be 100 km away, in a more abstract sense. Which of these should we consider to involve *perception* of spatial layout? This is an important question, because different neural mechanisms are likely to be involved depending on how one conceives of the term *perception.* The mountain example demonstrates that our expectations, prior experiences, and motivations (sometimes collectively called *top-down* factors) can play a role in shaping how we experience spatial layout, over and above what our senses tell us. While recognizing that the boundaries between concepts are likely to be fuzzy, many researchers define spatial layout perception in a way that largely excludes top-down influences. In this view, the *perceived* distance to the mountain would be 10 km. The neural pathways associated with top-down influences are poorly understood; because these influences involve complex comparisons between environmental features and information stored in long-term memory, the medial temporal and frontal lobes likely play a role. At any rate, the focus in this entry will be on the neural basis for spatial layout perception, considered separately from top-down factors.

Another important consideration is that spatial layout perception often carries the connotation that the perceiver is *consciously aware* of the layout of objects in the environment. This connotation is important because we may not always be aware of the spatial representations that the brain builds and uses. For example, the brain mechanisms that control eye movements, reaching, and other actions often rely on information about the spatial layout of objects in the environment, but these action-based spatial representations can come and go without necessarily influencing our awareness of spatial layout. This opens the possibility that the *consciously perceived* location of a coffee cup in front of us might be different than the internal coordinates the brain uses for the purpose of reaching to it. Many researchers do not consider these action-based representations to be perceptual in nature because they primarily exist outside conscious awareness.

Foundational Processes

Wakefulness

Based on the considerations previously discussed, it perhaps goes without saying that a person needs to be conscious in order to perceive spatial layout. Structures in the basal forebrain and thalamus (particularly the thalamic intralaminar nuclei and reticular nucleus) play crucial roles in regulating wakefulness.

Vestibular Processing

An important part of perceiving the layout of objects in the environment is determining how one's head and body are situated relative to the environment. Once the brain establishes some fundamental "cardinal directions" (such as which directions are up, down, and straight ahead), it can use this information as a reference when interpreting spatial information from the senses. For example, without knowing the orientation of one's head, it can be difficult to tell whether the head is vertical and the ground is sloping, or whether the ground is flat but the head is tilted. The vestibular system plays a vital role in establishing which way is "up" by responding to the force of gravity. Vestibular signals also interact with the visual system in determining the apparent orientation of the visual world with respect to the head: Dysfunction of the vestibular system can make physically vertical lines appear to be tilted, for example. The vestibular system also interacts with muscular signals in determining the apparent orientation of the body with respect to gravity. One source of evidence for this comes from patients with lesions in the posterolateral thalamus (a primary thalamic vestibular pathway). When these patients are sitting upright, they tend to perceive their body as tilted to one side. In an attempt to orient themselves so that they *feel* upright, they actively push their body in the opposite direction, making them unsteady and increasing the risk of falls. This is known as *pusher syndrome.*

The vestibular pathways run from acceleration-sensing organs in the semicircular canals of the inner ear, through the midbrain and thalamus, into several cortical regions. These regions include the temporoparietal junction, central sulcus, and intraparietal sulcus. Functional neuroimaging shows that the anterior cingulate cortex and hippocampus also tend to become active during vestibular stimulation. Relatively little is known about how vestibular signals are processed in the cortex. There is no "primary vestibular cortex" containing neurons that respond only to vestibular stimulation. Instead, neurons that respond to vestibular signals also typically respond to other forms of stimulation, such as vision or signals from the neck muscles. There is some suggestion that the right hemisphere plays a dominant role in vestibular processing.

Attention

Attention also plays a vital role in perceiving spatial layout. In the case of vision, attention guides our eye movements as we look at various objects in the world, and our perception of layout becomes progressively more fleshed-out in the process. The disorder known as hemispatial (or unilateral) neglect demonstrates not only the importance of attention in perceiving spatial layout but also provides insight into some of the neural mechanisms underlying spatial attention. Neglect patients often fail to respond to objects located on the side of their body opposite to their brain injury. For example, if asked to draw a daisy, neglect patients with a right hemisphere brain injury may draw petals and leaves on only the right side of the flower—"neglecting" the left side. The deficit is not due to blindness on the left side, because these patients do notice very salient objects or events in that region. A common interpretation is that this disorder represents a deficit in allocating attention to the left side. Although injury to several different brain areas can lead to neglect, it is most common after right hemisphere brain injuries involving the inferior portion of the posterior parietal lobe. This suggests that this region plays a role in allocating attention to objects and spatial layout on the left side of the body. There is some evidence that the analogous structures on the left side of the brain play a more general role, controlling spatial attention to both sides of the body.

Vision

Subcortical Pathways

One of the initial processing stages for visual spatial information occurs in the *lateral geniculate nucleus* (*LGN*) of the thalamus, a structure that receives input from the retina as well as feedback from the visual cortex. Even at this relatively early stage of vision, information about some object and spatial properties is kept physically separated from information about other properties. Information about color, depth, shape, and texture, for example, is processed in different layers in the LGN than information about object motion. After leaving the LGN, the fibers proceed along two neural pathways to the cortex. One path, called the *geniculotectal* pathway, carries visual signals from the LGN through the superior colliculus to cortical areas such as the posterior parietal lobe, largely bypassing the primary visual cortex. This pathway is important for controlling eye movements, but may play little or no role in conscious visual experience. One piece of evidence for this idea comes from the phenomenon of *blindsight*. In blindsight, individuals with brain injuries in the primary visual cortex experience "cortical blindness"—that is, they have no conscious experience of seeing anything at all in certain parts of their visual field. Nevertheless, they are sometimes able to point to objects in those regions with surprising accuracy. Although the interpretation of this accuracy has been controversial, some researchers see it as evidence of visual processing (without conscious awareness) in the intact geniculotectal pathway. The second pathway, called the *geniculostriate* pathway, primarily carries visual signals about color, depth, and form from the LGN to the primary visual cortex. This pathway is responsible for carrying the large majority of visual information used for perceiving spatial layout.

Early Cortical Processing

The first cortical way station for visual signals is the primary visual cortex, located toward the back of the head in the occipital lobe. Visual signals are processed to extract elementary visual features that form the initial basis for object recognition and spatial layout perception. Here, for example, the visual scene is decomposed into a collection of lines of varying widths and orientations. From this representation, the edges of objects are identified. Many cells here are *binocular,* in that they receive signals from both eyes. Importantly, some cells in the primary visual cortex are specialized for detecting *binocular disparity*—a strong source of binocular information about the relative depth between objects. Our eyes are located several centimeters apart in the head, and thus they each see the world from a slightly different perspective. This means that if we hold up two fingers at different distances from us, the images of the two fingers will be slightly closer together in one eye than they are in the other eye. This shift in the images between the two eyes is called binocular disparity; the greater the difference in depth, the greater the disparity. *Disparity-selective cells* are neurons in the primary visual cortex that receive signals from slightly different regions in the two eyes. Some cells respond best when an object appears in nearly the same region in both eyes (small disparity), while others respond best when an object appears in more widely separated regions in the two eyes (larger disparity). Because of the link between the amount of binocular disparity and the amount of depth between objects, these disparity-selective cells provide the brain with a strong source of information about spatial layout.

Beyond the Primary Visual Cortex

After leaving the primary visual cortex, visual signals undergo increasingly complex processing to detect more and more specific features of objects in the visual scene and their spatial layout. These neural signals follow multiple routes, the details of which are quite complex. For our purposes, however, cortical visual pathways may be summarized as following along two primary routes or streams: One stream runs dorsally, toward the top of the brain, from the primary visual cortex to the posterior parietal cortex (PPC). The other stream runs ventrally, along the lower part of the brain, from the primary visual cortex to the inferotemporal (IT) cortex. These two neural streams are thought to be specialized to perform different functions. This entry focuses on two particular frameworks for understanding this specialization.

"What" Versus "Where" Processing Streams

Leslie Ungerleider and Mortimer Mishkin proposed that the ventral stream encodes information about the *identity* of objects—that is, "what" objects are present, and the dorsal stream stores visual information about the *location* of objects relative to the observer—that is, "where" objects are. For example, when looking at a coffee cup, the dorsal stream processes information about the location of the cup, and the ventral stream recognizes that the object is a coffee cup. One source of evidence supporting this distinction came from studying the behavior of brain-lesioned monkeys. Monkeys with damage confined to the IT cortex were impaired at tasks requiring them to recognize objects by sight, but performed much better at tasks requiring them to appreciate the spatial relationship between objects. By contrast, other monkeys with damage confined to PPC were impaired at judging the spatial relations between objects, but performed normally in object recognition tasks.

"What" Versus "How" Processing Streams

More recently, David Milner and Melvin Goodale proposed a different account of the two visual streams. As in Ungerleider and Mishkin's earlier framework, they conceived the ventral stream as being specialized for encoding information about object identity ("what" objects are present). By contrast, they conceived the dorsal stream as being specialized for using vision to control actions to interact with objects. Thus, they characterized the function of the dorsal stream as not primarily representing object locations ("where"), but instead as representing "how" to move to interact with objects. One source of evidence supporting this "what" versus "how" framework came from studies involving a patient known by the initials "D. F." D. F. suffered damage to structures in the ventral stream pathway as a result of carbon monoxide poisoning. This event spared her intellectual abilities, language functions, and visual acuity, but left her with profound deficits specifically in using vision to recognize objects—a deficit termed *visual form agnosia.* She could readily name objects upon touching or hearing them, but had great difficulty naming them by sight alone. Most interestingly, she was able to reach out accurately to objects that she was unable to recognize visually. Other patients with damage in the dorsal stream, particularly the superior PPC, exhibit nearly the reverse pattern. They are able to visually recognize objects, but have difficulty reaching to objects while looking at them—a deficit termed *optic ataxia.* Taken together, these results support the idea that the functions of the dorsal and ventral streams may be understood in terms of using vision to control actions ("how") versus using vision to recognize objects ("what"), respectively, and that these functions operate separately to some degree.

Several features of this "what" versus "how" framework are particularly important for understanding the neural basis of spatial layout perception. First, in this framework, spatial information is thought to be processed in *both* the dorsal ("how") and ventral ("what") streams. Although the ventral stream plays an important role in object recognition, one must appreciate the spatial layout of the eyes, ears, nose, and mouth to be able to recognize a familiar face, for instance. Second, dorsal stream spatial representations are thought to be largely unconscious; they represent object locations in coordinates useful for controlling particular actions (such as eye movements or reaching), but are not closely associated with our conscious experience of spatial layout. Meanwhile, the *conscious* perception of objects and spatial layout is thought to be constructed in the ventral stream. Importantly, ventral stream representations can be used to control actions, particularly when an action is performed in the dark or must be conducted after a delay. This idea has been studied in healthy college students by using visual illusion figures to create errors in conscious perception. Some studies show that people reach or point accurately to parts of an illusion figure if they respond immediately, but their responses are more influenced by the illusion if they reach or point after a short delay. According to the "what" vs. "how" framework, the immediate actions are controlled by an accurate (and unconscious) dorsal stream representation of the figure's layout, whereas the delayed actions are controlled by an illusory conscious perception of layout constructed in the ventral stream.

Other Conceptualizations of Cortical Processing Streams

The "what/where" and "what/how" divisions of labor in the cortex are by no means the only frameworks that have been proposed. For example, Fred Previc has suggested that there are four fundamental behavioral systems for perceiving spatial layout. The *peripersonal system* uses vision to control reaching and grasping to objects very near the observer, and primarily involves the dorsal stream previously mentioned; the *focal extrapersonal system* controls visual search and object recognition, and primarily involves the ventral stream previously mentioned; the *action extrapersonal system* controls navigation and scene memory at greater distances, and primarily involves structures in the medial temporal lobe; finally, the *ambient extrapersonal system* controls spatial orientation, posture, and locomotion, and primarily involves subcortical vestibular pathways and signals for controlling muscle activity. This framework is representative of other current theories in one important way: It emphasizes that the distance of objects and the behaviors used to act upon them are often tightly intertwined—we reach to objects very near to us and walk to objects farther away. In many current theories, the brain mechanisms underlying perception of spatial layout are organized according to these intertwined factors. To some extent, these factors are further entwined with what coordinate system, or "frame of reference," the brain uses when processing the spatial information. For example, Sarah Creem-Regehr and Dennis Proffitt have proposed that the dorsal stream stores object locations in a format centered on the observer, whereas the ventral stream stores object locations relative to other objects in the environment.

Cortical Processing of Motion and Optic Flow

Relatively little is known about the neural basis of individual forms of layout information. In addition to binocular disparity, which is known to be processed in the primary visual cortex, another specific depth cue for which a neural basis has been identified is motion parallax. When riding in an automobile, a nearby guardrail appears to zip by rapidly; meanwhile, buildings on the horizon appear to move much more slowly. The brain can use this relative motion information to determine that the faster moving portions of the visual scene are closer to us. This cue, called *motion parallax,* not only helps us perceive the layout of objects around us as we move, but also their three-dimensional shape. The visual motion signals underlying this depth cue are processed primarily in a region near the junction of the occipital and temporal lobes known as the V5/MT complex, or MT+. In humans, functional neuroimaging studies have shown that moving patterns of dots (optic flow) activate this region. In addition, some individuals with brain injuries involving area MT+ have trouble perceiving visual motion—instead of seeing a fluid moving image, they see something akin to a series of stationary snapshots. Temporary inactivation of this region using powerful magnetic pulses can produce a similar effect, even in people without brain injury.

Putting It All Together

Signals from the dorsal and ventral streams converge on the hippocampus, a bilateral structure located deep in the medial temporal lobe (MTL). There is evidence that a nearby region in the MTL, namely the parahippocampal gyrus, becomes more active when people view pictures of places—that is, houses or other buildings—compared to when they view pictures of faces. This region has been dubbed the *parahippocampal place area.* In monkeys, some cells in the hippocampus become active when the monkey views a familiar scene from a particular vantage point. These view-selective neurons are called *view cells.* The MTL may be a region in which spatial layout perception is integrated with longer-term memories (top-down influences) about layout and landmarks.

As this entry makes clear, many brain areas must work together to support normal visual perception of spatial layout. Deficits in spatial layout perception can arise from injuries in a variety of brain regions. Damage to the primary visual cortex can preclude all visual layout perception by causing blindness; damage to MT+ can cause impaired motion perception. Bilateral damage to the posterior parietal cortex can cause *visual disorientation,* in which patients lose the ability to perceive spatial relations between objects. Other neurological conditions that may partly be due to impaired spatial layout perception are *topographical amnesia, topographical agnosia,* and *topographical disorientation.* Researchers have disagreed about how these terms should be defined, but each involves some kind of impairment in navigating. Typically, the deficits are interpreted as being due to difficulty in learning, remembering, or recognizing landmarks, rather than being due to visual perception of layout per se. Nevertheless, patients diagnosed with these conditions have rarely been tested systematically on their ability to perceive spatial layout, so layout perception may play some role. Damage to the posterior cingulate cortex, lingual gyrus, and parahippocampal gyrus have each been associated with difficulty in navigating.

Audition

The sense of hearing can provide important and compelling information about spatial layout. Sound can convey whether a room is a closet or a gymnasium. It can tell us when an unfamiliar person is walking behind us, or when a car is approaching at an intersection. A primary pathway for auditory signals runs from the cochlear nerve to the medial geniculate nucleus in the thalamus. From there, auditory signals continue on to the primary auditory cortex (PAC) in the superior temporal gyrus of the temporal lobe. In comparison with the functions of various regions of the visual cortex, relatively little is known about the function of auditory cortical regions. Interestingly, damage to the PAC does not cause deafness; the PAC apparently serves to represent complex acoustical properties of sounds rather than to process elementary auditory features such as pitch. At least some spatial characteristics of sound are processed in inferior portions of the parietal lobe—this region responds to moving sounds. There is also evidence that the right superior temporal cortex is engaged when sound sources change in distance.

John Philbeck

See also Auditory Processing: Central; Cortical Organization; Motion Perception: Physiological; Neuropsychology of Perception; Spatial Layout Perception: Psychophysical; Visual Processing: Extrastriate Cortex

Further Readings

Aguirre, G. K., & D'Esposito, M. (1999). Topographical disorientation: A synthesis and taxonomy. *Brain, 122,* 1613–1628.

Beckers, G., & Zeki, S. (1995). The consequences of inactivating areas V1 and V5 on visual motion perception. *Brain, 118,* 49–60.

Bisiach, E., & Vallar, G. (2001). Unilateral neglect in humans. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (2nd ed., Vol. 1). North-Holland, Amsterdam: Elsevier Science.

Brandt, T., & Dieterich, M. (1999). The vestibular cortex. Its locations, functions, and disorders. *Annals of the New York Academy of Science, 871,* 293–312.

Creem, S. H., & Proffitt, D. R. (2001). Defining the cortical visual systems: "What," "where," and "how." *Acta Psychologica, 107,* 43–68.

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The paraphippocampal place area: recognition, navigation, or encoding? *Neuron, 23,* 115–125.

Farrell, M. J. (1996). Topographical disorientation. *Neurocase, 2,* 509–520.

Fyhn, M., Molden, S., Witter, M. P., Moser, E. I., & Moser, M.-B. (2004). Spatial representation in the entorhinal cortex. *Science, 305,* 1258–1264.

Karnath, H.-O., Ferber, S., & Dichans, J. (2000). The origin of contraversive pushing: Evidence for a second graviceptive system in humans. *Neurology, 55,* 1298–1304.

Lamme, V. A. (2001). Blindsight: The role of feedforward and feedback corticocortical connections. *Acta Psychologica, 107,* 209–228.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action.* Oxford, UK: Oxford University Press.

Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, F., & Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nature Neuroscience, 3,* 1322–1328.

Previc, F. H. (1998). The neuropsychology of 3-D space. *Psychological Bulletin, 124,* 123–164.

Roe, A. W., Parker, A. J., Born, R. T., & DeAngelis, G. C. (2007). Disparity channels in early vision. *Journal of Neuroscience, 27,* 11820–11831.

Scott, S. K. (2005). Auditory processing—speech, space and auditory objects. *Current Opinion in Neurobiology, 15,* 197–201.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge: MIT Press.

SPATIAL LAYOUT PERCEPTION, PSYCHOPHYSICAL

Spatial layout in this entry refers to the three-dimensional arrangement of the surfaces that compose the visible environment of an observer. The visible spatial layout is typically a highly complex arrangement of surfaces. In an illuminated environment, light is reflected and refracted from the visible surfaces of the environment on its way to the eye. The array of light reaching the eye is thus complexly structured by the "optical projections" of the surfaces of the environment. Visual perception uses the structure of this "optic array" to obtain information about many properties of the spatial layout. This entry describes the information in the environment for the visual perception of orientation, content relations, location, scale and size, and distance.

The Spatial Framework

Humans are terrestrial animals, which means that they live mainly on the ground. The extended "ground plane" is the foundation for the perception of spatial layout. The floor of a room functions perceptually in a similar way to the ground plane. Other extended surfaces in the environment, such as the walls and ceiling of a room, join with the ground plane to form a visual framework surrounding the observer. Spatial layout is generally perceived in relation to this "spatial framework."

The spatial framework occupies much or all of the visual field, and the surfaces composing it tend to subtend large angles (called "visual angles") at the eye. Some important sources of visual information, such as linear perspective, are most perceptually salient at large visual angles. Also, some important perceptual effects of the spatial framework, such as the perception of being in a stable upright environment, depend on the large angular extent of the spatial framework.

Movement of the observer produces a continuous transformation of the entire optic array. There is a strong perceptual tendency to assume that such transformations are being produced by the observer's movement through an unmoving spatial framework rather than seeing the spatial framework as itself moving. This perceptual assumption is made even if the observer is not actively moving but is being passively carried through the environment (e.g., in a car or on a moving walkway). If an artificial visual environment is created that actually does move, while the observer remains stationary, the perceptual assumption of the stability of the spatial framework tends to produce the powerful perceptual illusion that it is the observer who is moving through a stationary environment.

Orientation

Many forms of visual information contribute to the perceived orientation of environmental surfaces, but some have a stronger effect on perception than others. Parallel receding lines on a surface, such as the floor of a room, project to the eye as lines converging toward a vanishing point on the horizon (Figure 1a). This is "linear perspective," which can produce a powerful and accurate perception of the three-dimensional orientation of the surface. If, however, the parallel receding lines are replaced by a textured surface with more irregular markings, the perception of recession is weaker and the amount of recession is perceptually underestimated, even if the same linear perspective is still implicit in the projection of this irregular texture. On the same receding surface, evenly spaced lines parallel to the horizon project to the eye as lines that are compressed more closely together with increasing distance (Figure 1a). This compression also contributes to the perception of the surface's orientation in depth, but it produces a weaker perceptual effect than linear perspective, and taken by itself produces much less perception of recession.

Because of their lateral separation, the left eye and the right eye have slightly different views of the world. This can contribute to the perception of surface orientation in depth through a perceptual

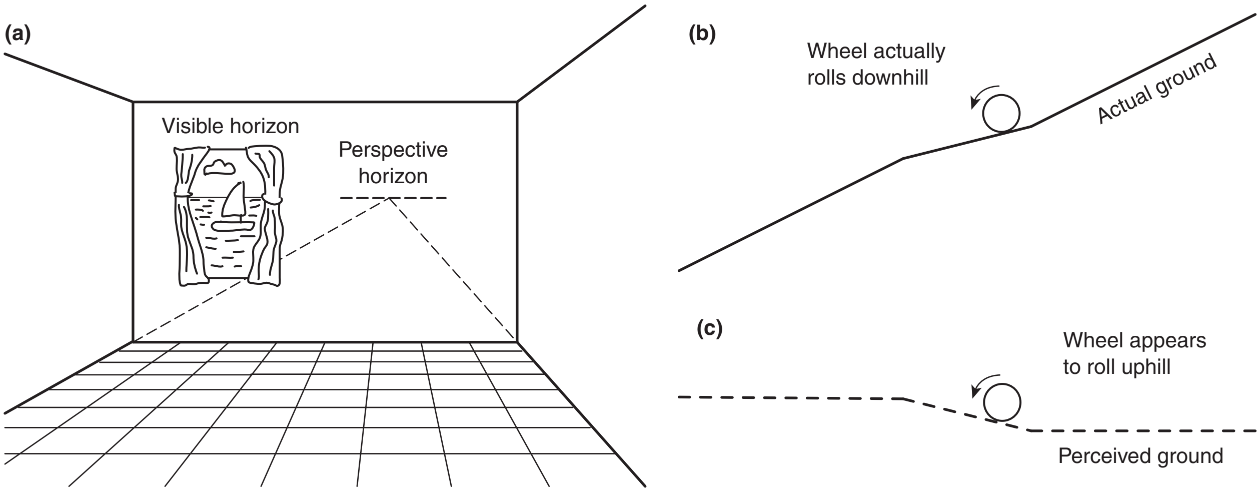


Figure 1 Orientation

process called "stereopsis." The differences between the two eyes' views are complex and depend on the orientation of the surface. For example, if the surface is slanted up from the horizontal, like a ramp, the two views mainly differ in their linear perspective, and this difference tends to produce a powerful and accurate perception of the surface orientation. If, however, the surface is slanted around a vertical axis, like a door, the two views mainly differ in their compression, and this difference tends to produce only a weak and inaccurate perception of the surface slant.

If an observer closes one eye and moves his or her head a few centimeters laterally, the open eye receives a changing view of the world, which is called "motion parallax." These changes in view are geometrically similar to the differences between the two views obtained with a stationary head by the left and right eye. Thus, it is not surprising that such head movements also contribute to the perception of surface orientation. As with stereoscopic slant perception, the effects of motion parallax are complex and depend on the orientation of the surface. If the changes are primarily related to linear perspective, they tend to produce a powerful and accurate perception of the orientation of a receding surface, but if they are primarily related to compression, they tend to produce a weaker and less accurate perception of surface slant in depth.

Surface orientation must be specified within some frame of reference. For example, if one says that a ramp is slanted 30° up from the horizontal, one is implicitly using an environment-centered frame of reference based on the horizontal ground plane and the vertical direction of gravity. There is ample evidence that the visible spatial framework has a powerful role in determining the frame of reference for the perception of surface orientation.

When an observer standing on a level ground plane looks toward the horizon, the sky fills the upper half of the visual field and the ground plane approximately fills the lower half. The horizon of a level ground plane is at eye level and so provides a visual reference for the horizontal orientation. Even if the horizon is hidden (by the walls of a room, for example), the surface features of the visible ground can specify the location of the horizon in the visual field. For instance, if the floor is tiled, the linear perspective produced by the tiles' receding edges will converge toward the horizon (Figure 1a).

The influence of the visible ground plane on the perceived orientation of other surfaces is made evident by powerful illusions that are produced when the orientation of the ground plane is misperceived. For example, sometimes one encounters landscapes where almost the entire visible extent of the ground plane is actually sloped downward (Figure 1b). In this situation, the ground plane will tend to be misperceived as horizontal, or nearly so, because it is itself the primary visual standard for the horizontal (Figure 1c). A local surface area whose downward slope is a bit less than that of the surrounding ground will then tend, by comparison, to be misperceived as sloping upward. Such localities sometimes

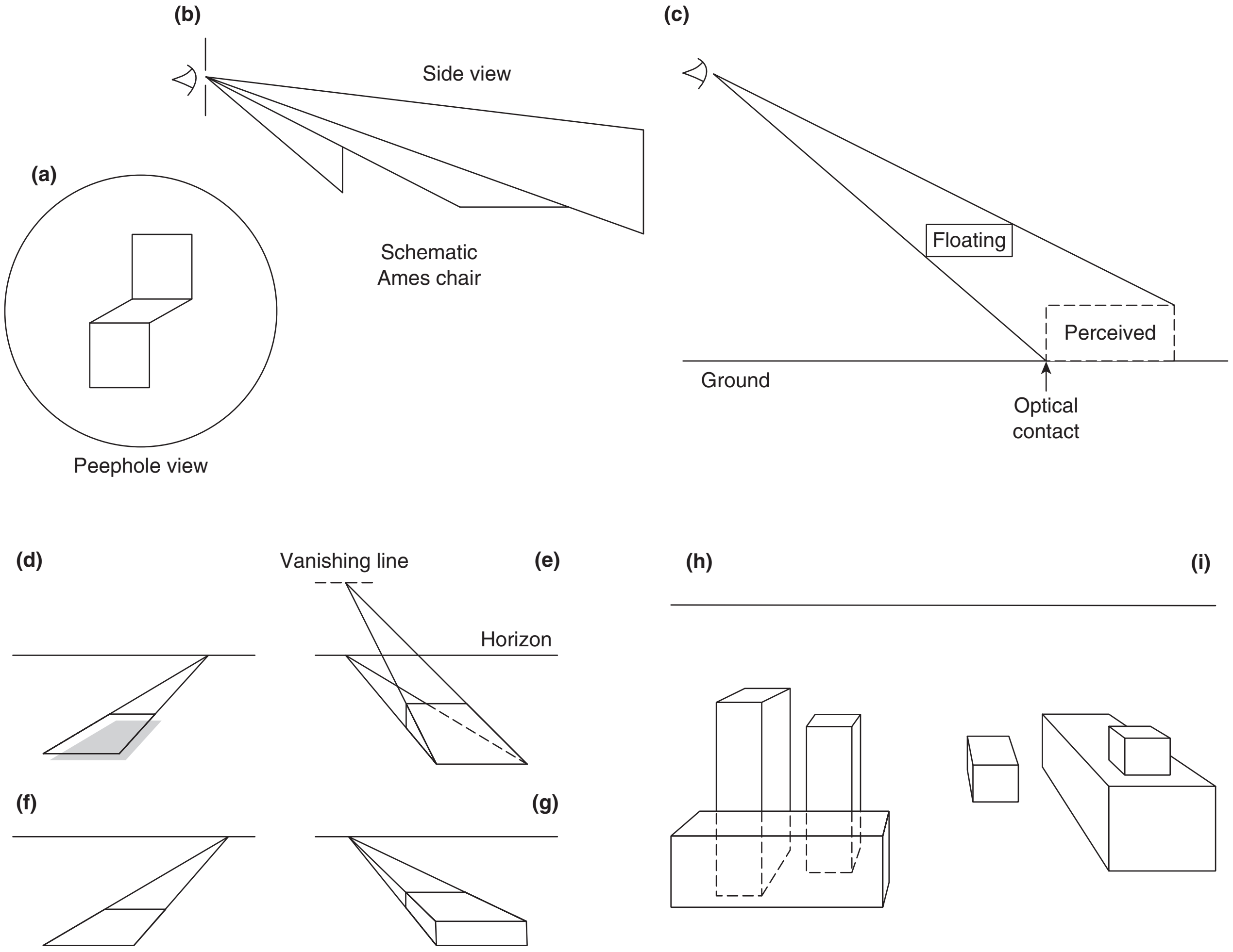


Figure 2 Contact Relations

become tourist attractions, "mystery spots" where the laws of physics are perceived to be violated. For example, a car that is actually rolling down the local slope (Figure 1b) will be erroneously perceived as rolling uphill (Figure 1c).

Contact Relations

Except for things that float or fly, everything is supported by the ground, either directly or indirectly. Thus, the surfaces that compose the visible environment are of necessity in contact with each other. The accurate perception of spatial layout involves the perception of these contact relations.

The visual perception of contact relations is not as straightforward as it might seem at first. If a visible gap separates one surface from another, then clearly they cannot be in direct physical contact with each other. But if the optical projection of one surface contacts the optical projection of another surface this "optical contact" does not necessarily imply physical contact. In the Ames chair illusion, for example, an observer looks through a peephole and sees what appears to be a chair made up of surfaces that are touching one another (Figure 2a); but in reality the surfaces are not physically touching (Figure 2b); they have been cleverly arranged so that from the viewpoint of the peephole they are in optical contact. This illusion illustrates the perceptual tendency to assume that optical contact implies physical contact. In this case, this tendency is easy to understand. From almost any other viewpoint there would be gaps or misalignments between the optical projections of the surfaces; the viewpoint from which their edges coincide exactly is a highly atypical one. There is a perceptual tendency to treat the observer's position as typical (or, at least, not highly atypical); this is sometimes called the "general position assumption." Many striking visual illusions, such as the Ames chair illusion, have been produced by violating this assumption, usually by requiring the observer's eye to be placed in a position that has been carefully chosen by the illusion's creator.

Not all contact illusions arise from forcing the eye into a highly atypical position. A box floating (or invisibly supported) above the ground will be in optical contact with the ground if the viewpoint is anywhere above the bottom of the box, and yet there will be a strong tendency to wrongly perceive the box as being in physical contact with the ground (Figure 2c). What is producing the illusion in this case could be called an "ecological support assumption," meaning that there is a perceptual tendency to perceive an object as resting on a supporting surface rather than floating above it (or in front of it, for example, in the case of a picture hanging on a wall). Various kinds of visual information can sometimes override this assumption by specifying that an object is above the ground rather than resting on it. If a flying carpet casts a shadow on the ground, for example, its shadow may produce the perception that it is hovering above the ground (Figure 2d). Similarly, movement of the observer's head will produce relative motion between the carpet and the ground, and this may also produce the perception that the carpet is hovering above the ground. However, the ecological support assumption is sufficiently strong that the carpet may instead be perceived to be sliding along the ground.

A complex set of geometrical constraints govern the possibility of contact between surfaces. For example, if two planar surfaces have different orientations, such as an inclined plane resting on the ground, then they can only be in contact along one edge (Figure 2e). A planar surface may be fully in contact with another surface if they are parallel to each other, like a carpet resting on the ground (Figure 2f). But parallel surfaces cannot be in contact if there are visible non-parallel links joining the surfaces, such as the sides that hold a top of a box above the ground (Figure 2g). Such geometrical constraints and their perceptual efficacy have not been extensively investigated, but the perceptual ability to correctly perceive the contact relations in many complex environments is evident. Sometimes, however, the spatial relations between surfaces are unclear because their contact relations are not visible (e.g., when the lower portions of objects are hidden, as in Figure 2h).

Rich natural environments may contain hierarchies of contact relations, such as a box lying on a larger box that in turn rests on the ground (Figure 2i). In such situations, a nested set of contact relations link surfaces to local supports, then to more global supports, and ultimately to the ground. When surfaces with different local supports have been linked to a common global support, they can then be spatially related to each other. Thus, for example, research has shown that observers are reasonably accurate in using perceived contact relations to match the distances of two boxes when one box is resting directly on the ground but the other box is lying on top of a table.

Location

Contact relations, when they are visible, establish an object's location within the spatial framework of the environment. For example, if a box is seen to be resting on a tiled floor, one also sees exactly which tile the box is resting on. An object's location can also be related in various ways to the locations of other objects. Some location relations are topological, such as an enclosure: If there is also a carpet resting on the floor, contact relations determine whether the box is resting on the carpet (and so is enclosed within its perimeter) or not (Figure 3a). Some location relations are ordinal: If several boxes are lined up on the floor, their contact relations with the ground specify their order (Figure 3a). These topological and ordinal aspects of location are fundamental but have not been studied much empirically, perhaps because they appear to be perceived easily and accurately when contact relations are accurately perceived. The perception of location also underlies the perception of the sizes and distances of objects within the spatial framework of the environment.

Scale and Size

The context provided by the spatial framework generally establishes a spatial scale at each location. The size of an object resting at a particular location may then be perceived in relation to the

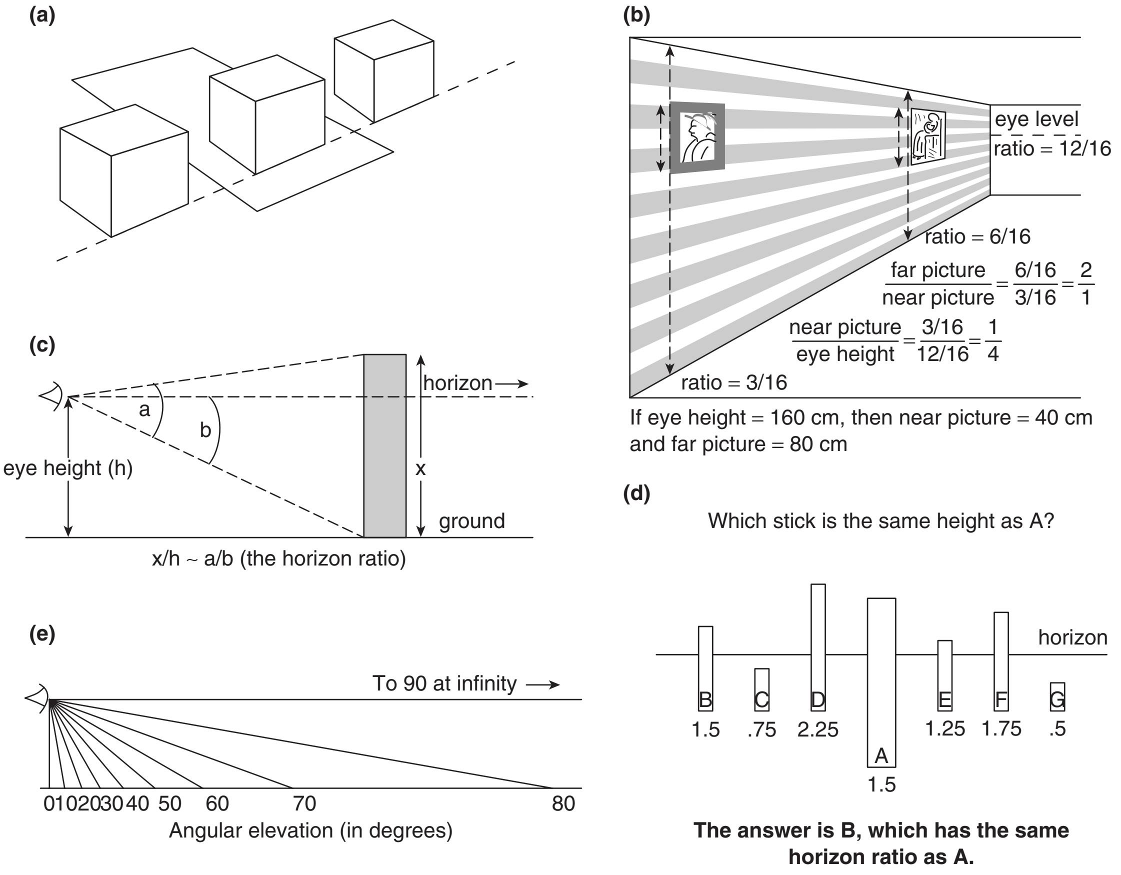


Figure 3 Location, Scale, and Distance

spatial scale at that location. One of the factors that contribute to establishing the spatial scale is simply the ratio between the angular extent of the object and the angular extent of its supporting surface. Consider, for example, the pictures hanging on the wall of a gallery. Even though the wall is receding in depth, so that its optical projection has a trapezoidal shape, the wall has a constant physical height along its entire length, and this establishes a constant vertical scale. The physical size of each picture is thus established by the ratio between its vertical angular extent and the corresponding vertical angular separation between the floor and ceiling of the wall (Figure 3b). A second scale factor is provided when a background surface has a visible local structure, such as a texture or pattern. If the units of local structure have a constant physical size (or range of sizes) across the extent of the surface, then they establish a "texture scale" for objects contacting the surface. The number of units covered by the object specifies its physical size. In the previous example of pictures hanging on the wall of a gallery, if the wall is striped, then the physical size of each picture is specified by the number of stripes it covers (Figure 3b).

A third scale factor is provided by the horizon of the ground plane. The line of sight to the horizon is parallel to the ground and so is at a constant physical height, equal to the height of the observer's eye, above the ground. Thus, for any object resting on the ground, the ratio between the physical height of the object and the physical height of the observer's eye is approximately equal to the “horizon ratio,” which is the ratio between the angular height of the object and the angular separation between the horizon and the location at which the object contacts the ground (Figure 3c). Analogous scale factors involving the horizon are potentially available when the observer moves or when the observer views the environment binocularly. However, their use in perception has not yet been investigated experimentally.

The perception of an object’s size can also be affected by the perception of its distance. For moderate to small visual angles, the relation between visual angle, distance, and physical size is approximately linear, so that the object’s approximate size can be obtained by multiplying its visual angle by its distance. Thus, distance can be thought of as a multiplicative scale factor. This scale factor of course depends upon the perception of distance (discussed in the following section).

When the observer’s location in the environment is visible, the scale factors previously discussed can generally be related to the observer’s own physical size. For example, the observer’s eye level on the wall of the gallery establishes the ratio between the observer’s eye height and the height of the wall, as shown in Figure 3(b). The scale factors then establish the physical sizes of objects relative to the size of the observer, as illustrated in Figure 3(b). This is sometimes referred to as “absolute size” because it allows size to be estimated in terms of some fixed unit of measurement, such as feet (historically derived from the size of a person’s foot), that is ecologically relevant to the observer. Research has shown that observers can make good use of such perceptual information in important practical tasks, such as climbing stairs or passing through an opening.

In some situations, however, the observer’s location in an environment is not clearly visible. For example, the observer may be looking out of an airplane window or may be looking at a photograph. In such situations, absolute size may not be perceptually available, but the “relative sizes” of objects may nevertheless be specified by the scale factors provided by the spatial framework. An observer looking at a photograph of a gallery may not be able to see whether it is a “life-size” gallery or a small model of a gallery, but the observer can still see, as in the previous example, that the farther picture on the wall is twice the vertical extent of the near picture (Figure 3b). Similarly, an observer looking at a photograph of sticks standing on the ground, and not knowing the height of the camera, can use horizon ratios to perceive which sticks are the same size. Figure 3(d) schematically illustrates an early experiment in which observers were able to accurately compare the sizes of such sticks out to great distances.

Distance

An object’s location within the spatial framework also provides a foundation for the perception of its distance from the observer and from other objects. The distance of the object is trigonometrically specified, relative to the eye height of the observer, by the angle of elevation of the observer’s line of sight to the object. However, this relation becomes highly nonlinear at large distances; each successive degree of increase in the angle of elevation corresponds to a greater increment of distance (Figure 3e). Although many studies have found that distance over the ground is usually perceived reasonably well, at large distances research has found perceptual underestimation that increases from less than 10% at around 15 meters (m) up to 25 to 40% at around 150 m; such underestimation may result from a perceptual failure to completely compensate for this nonlinearity.

Distances among different locations on the ground can also be perceived. Because of the incomplete compensation for the nonlinearity previously referred to, if one maps out the perceived distances among a set of locations scattered around on the ground plane, this perceptual map is somewhat compressed (by 15 to 50%) in its depth dimension relative to its lateral dimension.

Between any two locations in the environment there is a distance, but the general case of distance perception between locations on discontinuous surfaces has not been studied much. What evidence there is suggests decreased accuracy and increased variability in such situations. Even the perception of distance from the observer tends to be less accurate if the continuous ground plane between the observer and the object is visually obscured or disrupted by a gap or change in texture.

If an object’s contact relations with the spatial framework are not visible, then the object’s distance from the observer (called its “absolute distance”) can be difficult to perceive accurately. Nevertheless, some absolute distance perception is still possible, particularly at near distances. The accommodative adjustments of an eye's lens to bring an object into sharp focus vary with its distance, so in principle, perception could obtain distance information by monitoring these adjustments. Likewise, the closer an object is, the more the two eyes need to turn in to converge on it, so distance information could also be obtained by monitoring these vergence adjustments. Stereopsis could also contribute to absolute distance perception because the two eyes' views of a single object vary in ways that are related to the object's distance. For a variety of technical reasons, it is quite difficult to experimentally separate the contributions of these three sources of information for absolute distance perception, but many experiments have shown that when combined, they can produce a reasonably good perception of absolute distances within a meter or two of the observer.

At greater distances there is little perceptual information for the absolute distances of objects that cannot be visibly related to the spatial framework. When the observer moves, the angular change in visual direction of a single object decreases with its distance from the observer, but the evidence is weak that perception can use this "absolute motion parallax." As an object's absolute distance becomes very large, its appearance increasingly takes on the color of the surrounding atmosphere. This effect, known as "aerial perspective," has been used by landscape painters for many years to give an impression of distance, but it has not been studied much experimentally. It is evident that it can, at best, give a variable and crude impression of absolute distance.

Some theorists have suggested that anything can become a "cue" to perceived distance if it is consistently associated with distance in the experience of an observer, even if the association is arbitrarily created within an experimental laboratory. However, evidence supporting the generality of such an associative process in the perception of distance is quite limited. The known size of a familiar object, for example, has usually been found to have little effect on its perceived distance.

There is more visual information for the "relative distances" of objects from the observer. For example, if one object partially hides another object, this "occlusion" is perceptually compelling information that the occluded object is farther away. However, occlusion information is limited because it provides no information about how much farther away the more distant object is. Also, up to moderate distances, both stereopsis and motion parallax can support the accurate perception of the depth that separates two objects, even if these objects are seen in isolation from any spatial framework. Thus, for example, an observer using both eyes to view two luminous spots in a darkened room can use stereopsis to accurately adjust one spot so that it is at the same distance as the other.

Given how well stereopsis works without a spatial framework, it is remarkable how strongly a spatial framework can affect stereopsis. For example, if a luminous background is introduced behind two luminous spots, so that they are seen to be floating in front of it, then the background becomes the reference for their perceived depth. If the slant of the background is misperceived, the relative depth of the spots will be correspondingly misperceived. This occurs even though the relative depth of the spots is accurately perceived in the absence of the background. The influence of the background on stereoscopic depth perception is a striking example of the importance of the spatial framework in the perception of spatial layout.

H. A. Sedgwick

See also Binocular Vision and Stereopsis; Constancy; Direct Perception; Ecological Approach; Spatial Layout Perception: Neural

Further Readings

Epstein, W., & Rogers, S. (Eds.). (1995). *Perception of space and motion.* New York: Academic Press.

Gibson, J. J. (1950). *The perception of the visual world.* Boston: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Haijiang, Q., Saunders, J. A., Stone, R. W., & Backus, B. T. (2006). Demonstration of cue recruitment change in visual appearance by means of Pavlovian conditioning. *Proceedings of the National Academy of Sciences USA, 103,* 483–488.

Howard, I. P., & Rogers, B. J. (2008). *Seeing in depth* (2 vols.). New York: Oxford University Press.

Meng, J. C., & Sedgwick, H. A. (2002). Distance perception across spatial discontinuities. *Perception & Psychophysics, 64,* 1–14.

Ni, R., Braunstein, M. L., & Andersen, G. J. (2005). Distance perception from motion parallax and ground contact. *Visual Cognition, 12,* 1235–1254.

Sedgwick, H. A. (Ed.). (1986). Section IV: Space and motion perception. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1). New York: Wiley.

Sedgwick, H. A. (2001). Visual space perception. In E. B. Goldstein (Ed.), *Handbook of perception.* Oxford, UK: Blackwell.

SPATIAL MEMORY

A student walking to a classroom across campus must be able to stay oriented with respect to known locations within the campus environment. The consequences of becoming lost may only be a minor inconvenience to the student, but the consequences were surely much greater for our ancestors when navigating home from remote food sources. Current theories of navigation posit two types of spatial memories that work together to keep us oriented with respect to known locations. The first component is a long-term spatial memory of the environment through which the navigator is traveling. Long-term spatial memories contain distances and directions between objects, organized with respect to a spatial reference system centered on the environment. These long-term spatial memories play an especially important role in planning and executing navigational tasks.

The second component is a sensorimotor spatial memory, which contains distances and directions from the navigator to objects within the immediate environment. This memory is used to control actions such as walking around obstacles, passing through apertures (e.g., doorways), and moving toward intermediate landmarks. The sensorimotor spatial memory is limited in capacity, and therefore its contents are continually updated during self-movement; representations of objects are added and purged as a navigator moves from one locale to another.

Returning to the previous example, the student navigating across the campus must be able to identify salient features within the sensorimotor spatial memory (e.g., a recognizable building) and then match those features with the same features in the long-term spatial memory. If this matching process is successful, then the student can identify his or her location and orientation within the long-term spatial memory, thereby achieving an accurate sense of spatial orientation.

Perception plays an important role in both long-term and sensorimotor spatial memories. Whether learning from direct experience or indirectly through maps and other symbolic media, perception is the primary input used to generate long-term spatial memories of new spaces. Although most research on long-term spatial memory has focused on learning through vision, recent work indicates that many of the organizational characteristics of spatial memories are independent of learning modality. Additionally, perception of self-motion is critical to updating the body-to-object spatial relations contained within the sensorimotor spatial memory. Perception of body translations and rotations is based primarily on vision, proprioception, and vestibular stimulation. The rest of this entry covers some of the organizational principles of long-term and sensorimotor spatial memories, and address the roles of perception in both types of representations.

Long-Term Spatial Memory

Location is inherently relative and must be defined with respect to a spatial reference system. For example, your current location could be described in terms of your position within a room, your position within a city, or even in terms of latitude and longitude (assuming that you are located within a room, within a city, on Earth as you read this paragraph). These alternative definitions of your current location are all based on different spatial reference systems.

Similar to how cities on Earth's surface are defined with respect to latitude and longitude, locations in long-term spatial memory seem to be defined with respect to a small number of reference directions selected on the basis of cues, such as the shape of the environment and one's experiences within the environment. This reference direction organization of long-term spatial memory holds true across a wide variety of remembered environments, ranging in size from table tops to cityscapes, and ranging in realism from carefully designed laboratory environments to cluttered natural environments.

Perspective-taking performance is commonly used as an index of the reference directions used to represent a learned space. In a perspective-taking task, participants are asked to point to objects from imagined perspectives within a remembered environment (e.g., imagine standing in your kitchen at home, in front of and facing the sink, now point to the oven). Pointing responses are often made with a joystick, and pointing speed and accuracy are measured to assess the reference direction structure. Comparison of pointing speed and accuracy across different imagined perspectives reveals the underlying reference directions used to organize long-term spatial memory, as perspectives aligned with a reference direction are typically easier to imagine than are misaligned perspectives. This facilitation occurs because interobject spatial relationships aligned with a reference direction are stored in long-term spatial memory, whereas misaligned interobject relationships must be inferred. This inference process is cognitively effortful, resulting in increased pointing error and latency when imagining perspectives misaligned with a reference direction. Figure 1 depicts a hypothetical object layout. In this example, the layout structure (objects A–G), the rectangular room walls (solid lines surrounding the layout), and the location of the learning perspective (arrow below the layout) have resulted in a spatial memory organized around a single reference direction parallel to the learning perspective (indicated by the thick gray arrow). As a result of this organization, subsequent perspective taking is facilitated for perspectives aligned with the reference direction (e.g., imagine standing at C, facing D, now point to B), compared to misaligned perspectives (e.g., imagine standing at D, facing A, now point to B).

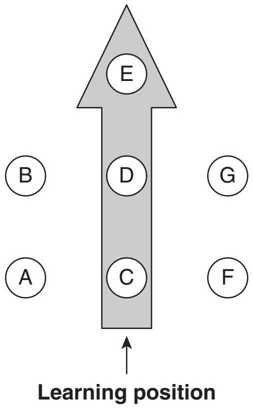


Figure 1 Example of Array of Objects With Assumed Reference Direction Indicated by Gray Arrow

The selection of reference directions occurs during learning and depends on two broad categories of cues. First, one's early experience of an environment, such as the first view, can be a cue to establish a reference direction parallel to the initially experienced view. This type of organization, based on the initially experienced view, can persist even after learning occurs from many perspectives. Second, the structure of the environment is a cue used to establish reference directions parallel to salient axes within the environment. For example, the walls of a rectangular room impose their own set of salient orthogonal axes parallel to the walls of the room. The organization of objects within the room, such as rows and columns of chairs in a classroom, can also have a similar effect. These salient environmental cues can sometimes override egocentric cues, such as the first view. One surprising result of this principle is that people can have better access to imagined perspectives that were never experienced, but are aligned with the environmental structure, than to experienced perspectives that are misaligned with the environmental structure.

Although environments are typically learned through vision, they can also be learned through touch, audition, and even language. Research on long-term spatial memories acquired through these nonvisual modalities is relatively sparse, but the emerging consensus is that these spatial memories reflect the same organization based on reference directions, no matter how they are acquired. However, the relative saliencies of cues for selecting reference directions may be quite different for different sensory inputs. For example, egocentric experience may be a more salient cue to selecting reference directions when object locations are learned through touch or sound because environmental cues, such as room shape, are more difficult to convey through these nonvisual sensory modalities. As such, the same layout may be represented quite differently depending on whether it was learned through vision or touch, but in both cases, the long-term spatial memory will be organized with respect to reference directions. The difference is in the specific reference directions selected.

The organization of long-term spatial memories in terms of reference directions may explain another important property of these memories, namely, their hierarchical organization. Memories of the locations of objects are organized categorically and hierarchically, such that a region of space may be represented as a whole, containing other regions and locations, and as a part, contained in larger regions. This property may result from the use of spatial reference systems at multiple scales. For instance, the spatial layout of each of the rooms within a house may be specified in a spatial reference system unique to each room. These spatial reference systems may serve as elements in a higher-order reference system defining the spatial relations among the rooms within the house.

The fact that spatial memories of environments learned through different sensory modalities adhere to similar organizational principles does not mean that perception is unimportant in spatial memory. It may be that perception actually underlies the reference frame organization. For example, after learning objects within a classroom filled with chairs arranged in rows and columns, long-term spatial memories of that scene will probably be organized with respect to reference directions parallel to the rows and columns of the chairs. In addition, eye movements made when learning this scene would be expected to follow the rows and columns of the chairs. This correlation between eye movements and selection of reference directions highlights the role of perception in creating long-term spatial memories. Gestalt grouping principles might guide eye movements along a specific sequence of objects, thereby influencing the reference directions selected during learning.

Sensorimotor Spatial Memory

Whereas long-term spatial memories are of great benefit when planning a route or reasoning about spatial relationships, they cannot be directly used to guide actions within the environment. Reaching for a coffee cup or walking toward a distant tree requires a representation of body-to-object spatial relations (referred to here as sensorimotor spatial memory), whereas the long-term memory only contains object-to-environment spatial relations. Because locations of objects with respect to the body in the sensorimotor memory change constantly as we walk and turn, this memory is not appropriate for long-term storage and is better understood as a working memory representation.

As with long-term spatial memory, the organizational properties of sensorimotor spatial memory can be revealed through perspective-taking performance. Whereas experiments on long-term spatial memory typically involve imagining perspectives within remote environments (such as imagining perspectives in one's home while seated in one's office), investigations of sensorimotor spatial memory involve imagining perspectives within the environment one currently occupies. Under these conditions, pointing responses made from imagined perspectives aligned with one's body are faster and more accurate than pointing responses made from perspectives misaligned with the body, indicating that object locations in the sensorimotor memory are represented egocentrically, in a body-defined framework. For example, it is easier to imagine facing the direction one is actually facing than to imagine facing the opposite direction. Recent accounts of this phenomenon suggest that the sensorimotor memory must be actively inhibited in order to imagine perspectives misaligned with the body. These misaligned perspectives are more difficult to imagine because this inhibition process is cognitively effortful.

Body-to-object spatial relations represented in the sensorimotor spatial memory must be updated continually during self-motion, a process known as *spatial updating*. Whereas reaching forward to pick up a coffee mug might be an appropriate action from one's current position, the same action would not be appropriate after turning 90° to the left or right. Instead, it is necessary to update the location of the mug with respect to the body as one turns to ensure the appropriate action from the new orientation. Research on the various perceptual cues to self-motion indicates that not all cues equally support spatial updating. These perceptual cues can be broadly categorized as idiothetic cues (internal cues, such as proprioception and vestibular stimulation) and allothetic cues (external cues, such as visual and auditory motion). Whereas idiothetic cues are often sufficient to perform spatial updating in the absence of allothetic cues, the reverse is not true. The importance of idiothetic cues is readily apparent when playing first-person video games, in which the user controls movement through the visual world by manipulating a joystick or a mouse. Such conditions provide allothetic but not idiothetic cues to self-motion and can rapidly cause the user to become lost in the video-game world. In contrast, adding idiothetic self-motion cues can make navigation much more natural. Virtual environments have proven to be an ideal tool for isolating idiothetic and allothetic cues to spatial updating, as participants can navigate through the exact same virtual environment by physically walking and turning or by manipulating a joystick. These experiments highlight the importance of idiothetic cues to spatial updating and the insufficiency of allothetic cues.

With sufficient self-motion cues, body-to-object spatial relations in sensorimotor spatial memory are updated continually when moving through the environment. Similar to findings on long-term spatial memory, where the reference direction organization is unaffected by learning modality, the body-based nature of sensorimotor spatial memory is also unaffected by learning modality. As such, imagined perspectives aligned with the body are facilitated for object layouts learned through vision, touch, audition, or even language.

Spatial Orientation

In order to stay oriented with respect to a known environment, the navigator must match represented features from the sensorimotor spatial memory with those same features in the long-term spatial memory. In some cases, this can be accomplished by matching identifiable landmarks, like the student who uses an identifiable building to stay oriented to campus. In other cases, geometric properties of the surrounding environment, like the shape of a rectangular room, can be used to perform this match. This matching process is a critical step to staying oriented to a remembered environment and underscores the importance of coordinating long-term and sensorimotor spatial memories.

Jonathan W. Kelly and Timothy P. McNamara

See also Action and Vision; Navigation Through Spatial Layout; Self-Motion Perception

Further Readings

Gallistel, C. R. (1990). *The organization of learning.* Cambridge: MIT Press.

May, M. (2004). Imaginal perspective switches in remembered environments: Transformation versus interference accounts. *Cognitive Psychology, 48,* 163–206.

McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology, 18,* 87–121.

McNamara, T. P. (2003). How are the locations of objects in the environment represented in memory? In C. Freksa, W. Brauer, C. Habel, K. F. Wender (Eds.), *Spatial cognition III: Routes and navigation, human memory and learning, spatial representation and spatial reasoning,* Lecture Notes in Artificial Intelligence (LNAI) 2685 (pp. 174–191). Berlin, Germany: Springer-Verlag.

Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B. (2004). Allocentric and egocentric updating of spatial memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 142–157.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*(6), 1157–1165.

SPEECH PERCEPTION

Speech perception refers to the processes involved in identifying and understanding the meaningful patterns of spoken language. The speech signal originates from the concerted actions of the speaker's lungs, larynx, jaw, tongue, lips, and soft palate (soft tissue in the back of the roof of the mouth) to generate sounds that are shaped in particular ways. A fundamental problem in speech perception is understanding how a listener recognizes the complex acoustic pattern of sound waves as being composed of meaningful linguistic units (vowels, consonants, syllables, words, sentences, etc.). This problem becomes strikingly apparent when one realizes that there is no simple one-to-one mapping between the acoustic speech signal and our perception of what the talker said. This entry examines attributes of the human voice and speech signal, some of the major experimental findings, and several prominent theories that attempt to shed light on the basic processes involved in speech perception.

The Speech Signal

The source of energy that drives the speech signal is the stream of air that originates from our lungs when we exhale. This air stream passes through our *vocal cords,* which cause the air stream to vibrate. When we talk, the vocal cords open and close rapidly, separating the air stream into a sequence of puffs of air. This sequence of puffs sounds like a "buzzing" noise, which changes in pitch as the cords vibrate faster or slower. The supralaryngal *vocal tract,* the part of the throat and mouth that lies above the vocal cords, further modifies the speech sound depending on its particular shape and size. Furthermore, by moving the soft palate, tongue, lips, and jaw (collectively referred to as *the articulators*), we can further alter the shape of the vocal tract and thus create a wide range of speech sounds.

One of the most basic characteristics of a human voice is its *fundamental frequency* (f_0), which corresponds to the perceived pitch of the speaker's voice (i.e., whether a person's voice sounds "deep" or "high"). Fundamental frequency is determined by the rate at which the vocal cords vibrate. Speakers have control over modifying the vibration rate while they talk, resulting in transient changes to f_0. These changes in f_0 can play a major role in the perception of different aspects of speech. For example, changes to f_0 can be used to emphasize one or more words (*word stress*). As another example, when an English utterance ends with a high pitch, this often signifies a question.

The articulators change the shape of the air stream and the frequency composition of the resulting speech waveform, which forms many of the common speech sounds of our language (e.g., vowels and consonants). The air stream can be wide open (resulting in vowels), redirected partially through the nose (resulting in the nasal consonants m and n), changed in shape over time (resulting in the gliding consonants j, w, and y), or momentarily stopped completely (resulting in the stop consonants b, d, g, p, k, and t). An important acoustical property of stop consonants is *voice onset time* (VOT), the delay between the burst of sound caused by quickly releasing a set of articulators (such as the lips) and the beginning of vocal fold vibration. For example, producing a syllable like "pah" requires the lips to release a burst of air to produce the /p/, and then a delay until the "ah" sound is made. "Pah" has a relatively long VOT. On the other hand, the syllable "bah," which involves more or less the same articulations, has a shorter VOT. As these examples illustrate, VOT can serve to distinguish speech sounds from one another. Even though the /b/ and /p/ are produced with the lips in similar ways before the vowel sound is made, it is the difference in VOT that distinguishes them.

Thus, speech sounds are generated through a complex combination of different vocal organs working together. Is it possible to identify basic units of speech from looking at the acoustic signal itself? Although most people are familiar with the idea of syllables and words in language, there is an even more elementary building block of speech: the *phoneme.* Phonemes are defined as the smallest units of sound that can distinguish one meaningful word from another. For example, consider the words *bat* and *bit.* These two words have identical sounds of /b/ and /t/ at the beginning and endings, but differ in the middle vowel sound. Thus, it is the middle elements—the phonemes (vowels in this case)—that distinguish these two words. Phonemes can be either consonants or vowels and can be further combined into larger units, including syllables, which generally consist of vowels surrounded by one or more consonants. Finally, groups of phonemes and syllables can be combined together to form meaningful words in a given language.

Despite the everyday assumption that speech can be broken into context-free discrete symbolic units, in reality the speech signal is not as tidy as the previous paragraph might lead one to believe. In fluent speech, the articulators begin to move into position to generate the next speech sound even while the current sound is still being produced. This property of speech production is called *coarticulation.* Coarticulation refers to the overlap that exists in speech production between the articulatory activity of adjacent phonemes. For example, the way in which you make a /k/ sound depends on the vowel that follows it. Because of the shape of the vocal tract necessary to produce different vowels, production of the consonant /k/

as in the word *key* requires placing the tongue farther forward in the mouth, compared to the /k/ in the word *coo,* where the tongue is placed farther back in the mouth. This results in differences in the acoustic speech signal for the two /k/ sounds, despite the fact that we perceive them as the same. These observations illustrate a general property of speech: The acoustic and articulatory features of a given phoneme are highly context dependent and are conditioned by the phonemes that precede and follow it. Strict context-free discrete perceptual units, such as vowels, consonants, phonemes, and even words, do not exist in the raw acoustic signal as they do in printed text. Rather, they are linguistic abstractions resulting from perceptual analysis. The speech sound is a continuous time-varying acoustic signal rather than a series of distinct units arranged sequentially in time. Thus, any given percept of speech may have many possible different ways of being represented at the physical level. Understanding the neural and cognitive processes involved in perceiving the abstract, idealized linguistic units of speech from the complex, context dependent, acoustic signal is one of the fundamental goals of speech perception research.

Some Important Findings

Early research in speech perception focused almost exclusively on the perception of phonemes in isolated syllables or nonsense words. One of the important findings observed in phoneme perception is called *categorical perception.* Categorical perception refers to the phenomenon of perceiving items from a large and varied stimulus set in terms of only a small number of discrete perceptual categories. For instance, consider a hypothetical example in vision. Suppose there is a wall painting colored red on the one end and yellow on the other, with the red color morphing very slowly and gradually into yellow, having no apparent distinct boundaries between the two. Generally, a viewer would look at the painting and notice the very smooth and continuous transition of colors, recognizing the myriad of color combinations between the two end points of red and yellow (e.g., "bright red," "reddish-orange"). But suppose a different viewer looked at this same painting and instead was convinced that she saw half of the painting as completely red and half as yellow, with no other colors in between. This is in essence the phenomenon of categorical perception: A physical stimulus that varies continuously is perceived as having a very clear-cut, well-defined, small number of categories.

Returning to the case of speech perception, consider the two syllables mentioned previously that differ only in terms of their VOT ("bah" and "pah"). Using modern speech synthesis techniques, it is possible to create artificial speech sounds of these two syllables and then continuously vary the VOT of these two stimuli to create a range of intermediate sounds. Using such techniques in the early 1950s, Alvin Liberman and colleagues at Haskins Laboratories found that when listeners are presented with these stimuli, they do not perceive the speech sounds continuously; rather, they identify either a clear "bah" or "pah" sound—with no intermediate percepts—even though the physical stimulus varies incrementally and gradually between the two sounds. It is as if speech is perceived like our second observer of the colored fields previously described. Thus, categorical perception may be a useful way for the brain to sort out a large, potentially confusing amount of variation in the speech signal into a limited number of more manageable, discrete perceptual categories.

Another important finding is that speech perception is significantly improved by visual access to the speaker's face. Especially under noisy listening conditions, being able to see the talker's mouth and articulators along with the acoustic speech signal significantly improves speech perception. For example, anecdotally, a common complaint among the elderly is that they are unable to understand what other people are saying unless they have their glasses on. However, it is not only under degraded auditory conditions that visual information has a functional impact on speech perception. A classic paper published by Harry McGurk and John MacDonald in the 1970s reported that auditory recordings of "ba" that were dubbed onto films of a person saying "ga" often led to reports of "da"—a "fused" utterance that was never actually presented. The *McGurk effect* is a multimodal perceptual illusion based on an unnatural co-occurrence of inputs. This illusion demonstrates that speech perception is susceptible to the influence of visual information even when the auditory signal is not degraded. This finding has led to interest recently in just how much information the visual channel can provide during speech perception, as well as what the underlying perceptual, cognitive, and neural mechanisms are that are involved in the integration of visual and auditory speech information.

The McGurk effect demonstrates how the acoustic signal of speech may be perceived differently given different visual contexts. However, similar phenomena have been shown to occur naturally in the auditory channel alone. For example, the existence of coarticulation effects demonstrates that a single acoustic cue can be perceived differently depending on other sounds in the immediate context. For instance, the same burst of sound can be perceived as a /p/ before a vowel such as "ee," but as a /k/ before the vowel "ah." Similarly, one's knowledge of the words of language can significantly influence the interpretation of spoken sounds. For example, when part of a familiar word is artificially obliterated with a patch of noise, listeners often report hearing the part of the word that was eliminated in addition to the noise. This is known as the *phoneme restoration effect.* Another example of this is that listeners often do not notice mispronunciations of sounds within highly familiar words. These findings illustrate how one's long-term memory and knowledge of language—"top-down" effects—can influence the interpretation of "bottom-up" perception of the acoustic speech signal. Thanks to the combination of top-down and bottom-up processes, our perceptual systems are remarkably robust and highly adaptable to both the wide variability that exists in the raw acoustic speech signal and the limitless range of contexts in which speech perception occurs.

The robustness of speech perception in the face of enormous physical variation gave rise to the traditional assumption that some kind of a *normalization* process occurs, with unnecessary redundant information being stripped away as a result of early perceptual analysis. That is, consistent with the findings of categorical perception, many researchers have implicitly assumed that the speech signal is reduced to an abstract, idealized linguistic message and that signal variability that is not directly related to the linguistic message is eliminated. According to the traditional abstractionist or symbol processing view of speech, perceiving speech is akin to perceiving printed letters on a page, with speech consisting of a linear sequence of discrete, idealized symbols.

Although much initial research was consistent with the idea that speech perception involved processes of normalization and abstraction, in recent years, it has become more apparent that much of the seemingly nonlinguistic information in the speech signal is retained and used to modify perception. For example, it is now known that so-called *indexical features* of speech—aspects of the speech signal that provide information regarding the speaker's identity and physical condition—can have an effect on speech perception. For example, indexical information in a particular person's voice may allow us to identify that the speaker is John, and that he is sad or tired. Recent studies have shown that indexical information in speech is not, in fact, discarded through a normalization process but instead may interact with memory and attention processes to affect how linguistic messages are perceived. As one example of how indexical information is used in speech perception, it has been found that familiarity with a talker's voice facilitates the accuracy of identification in noise of novel utterances spoken by that same talker. In sum, although this is a new direction of speech research compared to traditional methods of inquiry, it is now widely accepted that speech perception involves encoding both indexical information and the "symbolic" linguistic message, and that speech perception is necessarily influenced by both kinds of information in the signal.

Major Theoretical Perspectives

Several theories have been proposed over the years to explain the various phenomena and findings of speech perception. Three theories reviewed in this entry are the motor theory, direct realism, and the general auditory account of speech perception.

To begin, consider again the phoneme /k/ in *key* and *coo.* At the level of the acoustic signal, the /k/ sound is different for these two syllables, due to coarticulation effects involved in producing the two different vowels. However, listeners perceive both /k/ sounds as being perceptually equivalent, despite the differences in the actual speech signal. How listeners are able to perceive highly variable acoustic differences as equivalent sounds has been a major hurdle in our understanding of the processes of speech perception. The *motor theory of speech perception* (MTSP) attempts to bypass this hurdle by proposing that listeners unconsciously articulate the speech sounds they hear and then use their own articulation to perceive and understand what they heard. The reason we perceive equivalence despite the underlying physical differences in the speech signal, proponents of MTSP argue, is that it is the same essential articulatory gesture or motor command that is used to produce /k/ in both cases. That is, our brains attempt to map what we hear onto something that it knows how to produce, using its motor categories to define what is perceived.

If this version of MTSP is correct, then the underlying articulations producing speech ought to be less variable than the actual acoustic signal. However, it was soon discovered that low-level articulatory motor activity is no less variable than the actual acoustic signal. For example, as already discussed with *key* and *coo,* there are slight differences in how we produce the /k/ sound. Both within and across individual speakers, articulatory variability is high, even when the perceptual result is relatively stable. Motor theory was subsequently revised, such that the proposed motor correspondence no longer referred to externally measurable articulatory motions, but rather to the recovery of abstract sets of motor commands. Some recent evidence has shown that the brain contains *mirror neurons*—neurons that are active both when a person produces a particular action and when the person observes someone else producing that same action—which could provide a neural account of MTSP. Even so, with the move away from external articulations to internal motor commands, MTSP's main hypothesis has become extremely difficult to test. Another problem with MTSP is that it was found that chinchillas appear to show categorical-like perception for human speech sounds, despite the inability to produce speech themselves. Although MTSP may be less tenable than originally thought, researchers continue to explore the idea that speech production and speech perception are closely linked.

The *direct realist* (DR) approach to speech perception is based on the legacy of James Gibson's ideas of "direct perception," a theory of perception that argues that the senses provide us with direct awareness of the external world, rather than having our perceptions based on some internal representations of the world. Direct realism is similar to MTSP in that there is an emphasis on articulatory events rather than the acoustic signal. However, rather than relying on an internal, abstract motor command, proponents of this view argue that there is no need to examine the internal contents of the perceiver to explain speech perception. In other words, there is no need for positing intermediate perceptual or cognitive mental representations. Direct realism also differs from MTSP by proposing that speech perception involves domain-general mechanisms of perception that are also used in nonspeech domains, such as vision, whereas MTSP argues for speech-specific perceptual mechanisms. The DR perspective provides a valuable reminder of the need to step back and consider speech perception in relation to the larger environment. The DR approach has also been an important framework that has led to a better understanding of the sources of information that are available in the acoustic signal itself. However, it has been difficult to design new methods to test the basic claims of DR. Furthermore, the DR view runs counter to the current mainstream perspective in the psychological sciences that emphasizes perception and cognition as consisting of stages of information processing and the manipulation of internal representations. Possibly for these reasons, DR currently represents a minority view in speech perception research.

Finally, the *general auditory account* (GAA) proposes that speech perception can be explained by general-purpose mechanisms and processes that are common to audition more generally, not just specific to speech perception. In this way, GAA is similar to DR. However, GAA differs from both of the previous two theories in that it assumes that speech perception relies on the acoustic signal itself rather than on the perceiver's underlying motor or articulatory gestures. How does GAA explain perceptual equivalence of speech sounds despite the large amount of underlying variability in the speech signal? The proposal is that perceptual equivalence is due to a general ability of the perceiver to learn and make use of multiple acoustic cues or sources of information in the signal to narrow in on a single perceptual category. Perceptual equivalence thus arises from the integration of multiple cues in the speech signal, where any single cue alone is imperfect, but the combination of many cues together can be predictive and reliable. The advantage of GAA is that it does not rely on specialized speech-specific mechanisms in order to explain the basic phenomena in speech perception; thus, findings from any perceptual domain, including from nonhuman animals, can provide theoretical insight. However, a possible disadvantage of GAA at this time is that it is considered to be vague and unspecified, rather than being a coherent theory, and thus presently may be limited in the explanatory power it can currently provide relative to other theories.

Several prominent theories have been proposed to explain speech perception, but it appears that no single theory at present can adequately explain all findings.

Christopher M. Conway, Jeremy L. Loebach, and David B. Pisoni

See also Audition; Computer Speech Perception; Perceptual Development: Speech Perception; Speech Perception: Physiological; Speech Production; Word Recognition

Further Readings

Cleary, M., & Pisoni, D. B. (2001). Speech perception and spoken word recognition: Research and theory. In E. B. Goldstein (Ed.), *Blackwell handbook of perception* (pp. 499–534). Malden, MA: Blackwell.

Denes, P. B., & Pinson, E N. (2007). *The speech chain: The physics and biology of spoken language.* New York: W. H. Freeman.

Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech perception. *Annual Review of Psychology, 55,* 149–179.

SPEECH PERCEPTION: PHYSIOLOGICAL

Speech, as our primary means of communication, is perhaps the most important sound in our daily lives. Current theories of how the brain perceives speech rely on more than a century of investigation that has included patients with brain damage, microelectrode recordings in nonhuman animals, electric fields measured on the human scalp, and, most recently, neuroimaging. The current consensus, as described in this entry, holds that speech proceeds along parallel pathways or streams in the cortex. A “what” stream is dedicated to speech comprehension, and a “where” or “how” stream is more important for learning speech and holding it in mind (as when you remember a phone number in your head; color insert, Figure 6). The left hemisphere of the brain tends to be dominant for many aspects of speech perception, such as understanding sentences, though the reasons why are actively debated.

Neural Measures

Numerous techniques have been used to study how the brain processes speech. One of the oldest is lesion analysis, where functional neuroanatomy is inferred from patients with localized brain damage (lesions, e.g., from stroke or trauma) who exhibit a particular language deficit or aphasia. Until the middle part of the 20th century, much of our understanding of speech and the brain came from lesions. With the widespread use of microelectrode recordings in nonhuman animals, researchers began to characterize moment-by-moment representations of sounds. Microelectrode recordings, though, are surgically invasive and cannot be performed in healthy people. Most early studies of real-time speech perception in humans instead used electroencephalography (EEG), which measures electrical fields from neural activity with electrodes resting on the scalp. Numerous characteristic deviations or oscillations in electrical waves have been identified in speech and language processing. However, early EEG studies did not use many electrodes on the scalp and could not identify *where* in the brain this speech processing occurs.

Recent decades have witnessed a flood of speech studies using neuroimaging, which aims to localize brain function. Functional magnetic resonance imaging (fMRI) for instance, developed in the 1990s, measures changes in the blood supply to infer where and roughly when (within about a second) neural activity occurs. High-density EEG and magnetoencephalography (MEG) use many (typically more than 100) electrodes or magnetic sensors to measure neural activity. Like the traditional EEG, they have excellent temporal resolution, but they can also estimate roughly where in the brain activity occurs. Due to their relative strengths, fMRI is now the dominant technique for studying localization of speech processing, and EEG/MEG are the dominant techniques for measuring its timing.

Cortical Organization

The classical understanding of the speech and language cortex arose in the late 19th century, when two physicians, Paul Broca, and Carl Wernicke, linked different aphasias to specific lesions in the brain. A part of the superior (top) and posterior (back) temporal lobe (the part of the brain just above your ears) came to be known as *Wernicke's area,* which when damaged caused patients difficulty in comprehending speech. Part of the left inferior frontal lobe came to be known as *Broca's area,* which when damaged caused deficits in speaking fluently. The natural conclusion was that Wernicke's area handles perception and Broca's area handles production. In recent decades, however, researchers have grown to appreciate the limitations and complexities of this classical view.

For instance, it turns out that damage to Broca's area may have less impact on fluency than lesions of a deeper-lying structure called the *insula,* which is not traditionally a core member of the speech pathway. Furthermore, damage to Broca's area is characterized not only by effortful, nongrammatical speech, but also difficulty understanding grammatically complex speech. Wernicke's aphasics, on the other hand, may have poor speech comprehension but may also produce speech that—while fluent—is devoid of meaning or contains many errors. Furthermore, Wernicke's area has been redefined anatomically in many different ways and has at least several functionally distinguishable parts. Some of the goals in contemporary research are to explain these complexities and to label brain areas and networks by the computations they perform, rather than the behavioral syndromes their damage causes.

There is no single, universally accepted view of the speech and language cortex, and a lively debate persists. However, there is a growing consensus that speech and other acoustic information follows at least two parallel streams of processing from the early auditory cortex: a ventral (lower) and a dorsal (upper) stream. The ventral stream processes speech for comprehension. This "what" pathway projects from the early auditory cortex, located on top of the temporal lobe, down the side of the superior temporal lobe. It appears to project both forward and backward along the superior temporal lobe, though the functional importance of this divided ventral stream remains unclear. Some investigators believe that the fundamental sound-to-meaning mapping of the ventral pathway occurs in the posterior tributary, where acoustic word forms (e.g. the sound "cat") are matched with their widely distributed conceptual representations (the thought of a furry pet). On this view, the anterior tributary is devoted only to higher "combinatorial" aspects of language, such as processing syntax in complex sentences.

Other researchers hold an equally well-supported view that emphasizes a hierarchy of areas along the anterior tributary. Along this anterior superior temporal pathway, neural preferences develop systematically from simple sounds (not just speech), to speechlike sounds and the constituents of speech (phonemes), to wordlike sounds, to intelligible words and sentences. At its anterior end, neurons in this tributary have abstracted away the acoustics of speech and reflect intelligibility irrespective of the speech's physical attributes. Whatever their distinguishing roles, though, both anterior and posterior tributaries of the ventral stream appear to be important for normal speech comprehension.

The dorsal stream projects back and up from the superior temporal lobe into the inferior parietal lobe (the top, back part of the brain) and then to the inferior frontal lobe, including Broca's area. This stream is thought to process speech more for articulation, as when learning language or new vocabulary, and maintaining transient representations of the speech, as when remembering a phone number by saying it in your head. The dorsal stream is also important for localizing sounds in space. Historically, the ventral and dorsal streams were introduced in the visual system simply as "what" and "where" pathways, for recognizing and localizing objects respectively. However, in both audition and vision, dorsal stream descriptions have generalized to include sensorimotor integration, or perception for action as opposed to comprehension—so, not just "where" but "where and how." This fits well with a long-held theory that speech perception relies intrinsically on articulatory representations (the *motor theory*). It may be that there are separable "where" (sound location) and "how" (articulation, short-term memory) subsystems in the dorsal stream, or they may both perform qualitatively similar sensorimotor operations, but with different systems for action (eyes versus mouth and tongue).

Hemispheric Dominance

Speech is one of the clearest examples of how the two hemispheres of the brain can show dominance for different functions. In most people, the left hemisphere is dominant for the bulk of speech and language processing, both comprehension and especially production. Many lines of evidence support this lateralization.

Behaviorally, most people show a *right ear advantage* (REA) for speech comprehension. Because the strongest neural connections cross the midline between the ear and the cortex (right ear projects more strongly to left hemisphere and vice versa), the REA implies a *left* hemisphere advantage. If a person listens simultaneously to different word lists in each ear and then reports which words they heard, they typically report many more words presented to the right ear (i.e., the left hemisphere) rather than the left ear. Even clearer evidence for left-hemisphere dominance comes from the field of neurosurgery. Before surgery, patients may be evaluated for hemispheric dominance with the Wada or sodium amytal test, in which one hemisphere is anesthetized while the patient is tested for speech function. Most people lose the ability to communicate verbally when their left hemisphere is anesthetized but not their right. Interestingly, handedness is well correlated with hemispheric dominance as measured by the Wada test: About 97% of right-handers and 75% of left-handers (who only make up about 10% of the population) are left-hemisphere dominant.

Another patient group, with so-called split brains, provides a slightly more subtle understanding. In split-brain patients, the connections between the hemispheres are surgically severed to control epilepsy. This results in two somewhat independent brains, both apparently aware, that can be tested for their speech and language ability. The left hemisphere in a split-brain patient can understand natural language with complex syntax and can respond verbally. The right hemisphere seems capable of understanding words but cannot use syntax and can only respond by pointing or gesturing. These results, along with many lesion and neuroimaging studies, argue that laterality is not exclusive for all stages of processing. The early stages of speech comprehension may be more bilateral (though perhaps with a leftward bias), compared to later stages of comprehension and production, and the "what" stream may be more bilateral than the "where/how" stream.

Although the observation of left-hemisphere dominance is uncontroversial, different explanations have been proposed. One influential view holds that the left hemisphere is better at processing fast spectrotemporal acoustics, and the right hemisphere is better at slow ones. The rationale is that the features of speech most informative for content often change rapidly through time (in English, anyway), so left laterality is a consequence of the left-hemisphere's advantage for fast temporal processing. The right hemisphere instead is thought to be better at processing pitch or slowly changing features as in prosody (the rhythm, stress, and intonation of speech). Many studies support a right-hemisphere preference for these attributes, for example, for "voice" or speaker identity processing in the right temporal lobe. Moreover, some neuroanatomical asymmetries are consistent with this view, such as greater myelination (the insulation on neural fibers, which makes for faster impulses) in the left auditory cortex. However, not all studies find a neat hemispheric dissociation between fast and slow processing.

Another explanation for speech laterality is that the hemispheres differ in the temporal scales over which they integrate information. This *multi-time resolution* theory holds that the right hemisphere preferentially integrates information over longer time windows (150–300 milliseconds), whereas the integration over shorter timescales (20–50 milliseconds) occurs more bilaterally. Finally, a more recent suggestion is that the left hemisphere deals with more categorical representations, which happen to be more appropriate for speech sounds and words. This explanation is interesting because it puts the functional distinction between hemispheres at a more abstract level and may accommodate a broader range of evidence than theories emphasizing time.

Lee M. Miller

See also Aphasias; Audition: Cognitive Influences; Auditory Processing: Central; Auditory Scene Analysis; Auditory System: Structure; Speech Perception; Speech Production; Speechreading

Further Readings

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience, 8*(5), 393–402.

Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., & Medler, D. A. (2005). Neural substrates of phonemic perception. *Cerebral Cortex, 15*(10), 1621–1631.

Moore, B., Tyler, L., & Marslen-Wilson, W. (2008). The perception of speech: From sound to meaning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 363*(1493), 917–921.

Price, C., Thierry, G., & Griffiths, T. (2005). Speech-specific auditory processing: Where is it? *Trends in Cognitive Sciences, 9*(6), 271–276.

Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences, 26*(2), 100–107.

SPEECH PRODUCTION

Speech is produced by joint activation of three subsystems—respiratory, phonatory, and articulatory (see Figure 1). The respiratory system produces an outflow of air under relatively constant pressure. The modulation of this airflow to produce the sounds we know of as speech is characterized by the source-filter theory, according to which the phonatory system provides the sound source for speech and the articulatory system filters this source.

The phonatory system acts on the outward airflow of the respiratory system to produce phonation, or voicing. The sound source is the vibration of the vocal cords (or folds), which are two opposed horizontal muscular slivers running from front to back within the cartilaginous larynx. They are drawn apart during quiet breathing. Prior to the first cycle of vocal fold vibration for speech, the vocal folds are brought together. This blocks the airway causing air pressure to build up beneath the folds. The increasing pressure evokes an elliptical opening between the folds through which a puff of air flows. But the opening evokes two counter-effects that restore the closed state. They are the elastic recoil of the stretched folds and the Bernoulli effect—the development of low pressure in regions of high flow—which sucks the folds together. The folds close immediately, pressure builds to eventually blow them apart again, and the next cycle continues. In an adult male phonating a vowel, the rate of vibration is about 125 hertz (Hz) (cycles per second). The phonatory system also produces a basic distinction between voiced consonants (during which the vocal folds are vibrating), and voiceless consonants (during which they do not vibrate). Vibration is eliminated by moving the folds laterally (to the sides) into the open breathing position. In this position, they are not activated by the air flow.

The successive puffs of air generated by vocal fold vibration enter the supralaryngeal vocal tract—the air space between the larynx and the lips. They not only produce vibration that propagates through the air to the listener at their repetition rate, called the fundamental frequency (f_0), but they also produce vibration at the first 40 or so whole numbered multiples (harmonics) of f_0.

The main role of the articulatory system is to modulate the acoustic package (collectively known as the source spectrum), which enters the supralaryngeal vocal tract. The supralaryngeal vocal tract (oral and pharyngeal cavities) acts as a filter. It provides a set of variable resonators selectively amplifying different parts of the source spectrum for vowels and voiced consonants depending on the size and shape of cavities formed in the tract.

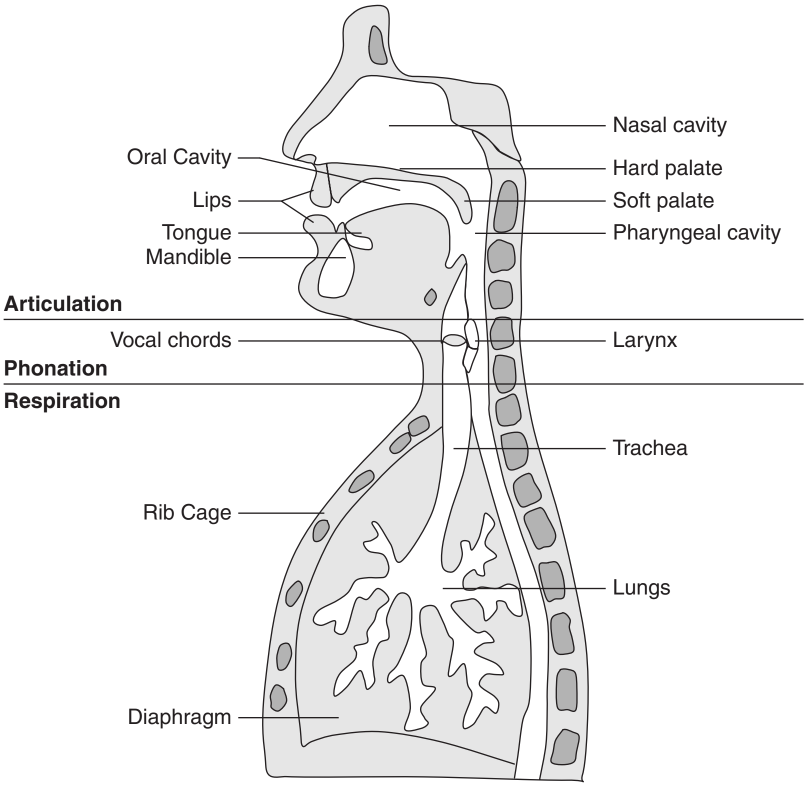


Figure 1 Sketch of the Three Subcomponents of the Speech Production Apparatus

This is analogous to how the tuning knob on a hi-fi stereo selectively amplifies a certain small subset of radio frequencies in each individual position that one turns it to. In the simplest speech-related case of the so-called neutral vowel (e.g., the first vowel in the phrase "the book"), the tract assumes the shape of a single tube of uniform cross-sectional area. In a typical male adult, this configuration maximally amplifies frequencies of 500, 1500, and 2500 Hz. In general terms, these frequencies are known as *resonant frequencies,* but in the specific context of speech they are called *formants*. For other vowels, and voiced consonants, involving a nonuniform tract, the oral and pharyngeal cavities tend to make separate contributions, with the pharyngeal cavity responsible for the first (lowest) formant, and the oral cavity responsible for the second. Topics discussed in this entry include the syllable, individual consonants and vowels, suprasegmental phenomena, and Tinbergen's four questions.

The Syllable

The serial organization of speech centers on an alternation between a relatively closed mouth for consonants, producing a constriction in the supralaryngeal vocal tract, and a relatively open (and unconstricted) mouth for vowels, produced by oscillation of the lower jaw or mandible. One iteration of this cycle comprises a consonant-vowel syllable, the only syllable type present in all languages. More generally, a syllable can involve anything from a vowel alone (as in the "a" in *about*) to a vowel with several consonants on each side (as in *strengths*).

Individual Consonants and Vowels

Consonants are traditionally classified in terms of three basic properties. The first property is the place of articulation. This refers to the place in the vocal tract at which the constriction is made. For example, in English "p" and "b," the constriction is made at the lips, whereas for "t" and "d," it is made by the tongue just behind the upper teeth. The second property is the manner of articulation. This refers primarily to how much of a constriction is made. For example, for the previous sounds, called stop consonants, there is a total occlusion of the tract. In contrast, for fricatives such as "s" and "z," air flows turbulently through a small aperture. The third property is voicing. This refers to the extent to which the vocal folds are vibrating during the sound. For the voiceless sounds "p," "t," and "s" they do not vibrate at all. For forms of "b" and "d," voicing is more characteristic than for "p" and "t," and "z" is voiced throughout the sound. Vowels are characteristically classified in terms of the height of the tongue and its position in the front-back dimension.

Suprasegmental Phenomena

In addition to the syllable, there are two other phenomena that transcend individual consonants and vowels, and are therefore called suprasegmental. Many languages have stress systems that involve differences in emphasis on syllables. For example, in the word *syllable,* the first syllable is stressed. Another suprasegmental phenomenon is intonation, which is a pattern of fundamental frequency variation across an utterance, heard as pitch variation. For example, in a statement such as "Joe ate his artichokes," the fundamental frequency goes down at the end of the utterance, whereas for a question such as "Did Joe eat his artichokes?" fundamental frequency goes up at the end.

Tinbergen's Four Questions

A brief conception of speech production from the evolutionally oriented perspective of ethology (the study of naturally occurring behavior), including some implications for speech perception, can be obtained by addressing four questions that Nobel Laureate Nikko Tinbergen posed for any communication system.

1. How Does It Work?

This entry has already considered what happens at the periphery, but how is production centrally controlled? The occurrence of speech errors in which individual segmental units (consonants and vowels) are misplaced in an otherwise correct utterance tells us that these units and their serial organization are separately controlled. This organization can be best described in terms of a frame/content metaphor. The frame is formed by the basic alternation between a closed and open mouth, which characterizes the syllable. The content consists of the particular consonants and vowels that are inserted into the frame. The existence of separate frame and content components is revealed by the fact that misplaced vowels go into vowel positions and misplaced consonants go into consonant positions. For example, in a vowel error, "high heels" turns into "he hiles," and, in a consonant error "better form" turns into "fetter borm." In contrast, consonants and vowels do not occupy each other's positions in errors, which would, for example, allow "abstract" to turn into "bastract." Thus, there is a separate specification of the positions in the sequence into which vowels and consonants must be inserted—a frame specification—and a particular sequence for consonants and vowels themselves—the content. Underlying this dual output control process are two brain subsystems, usually in the left hemisphere of the brain: a lateral system, including the classical Wernicke's and Broca's areas, primarily concerned with content, and a medial system including the premotor cortex of the supplementary motor area, primarily concerned with frames.

2. What Is Its Functional (Adaptive) Significance?

Speech allows the symbolic transmission of an extraordinarily large number of words in the form of individual pieces of symbolic vocal substance, each attached to an individual concept (e.g., the sound complex "dog" attached to the concept of the domestic canine in English). Many think that the advent of this concept-symbol pairing capacity could have been the single most important event in human evolution. What made large vocabularies possible is the emergence of a particulate structure whereby a limited number of particles—consonants and vowels—packaged into slightly larger particles—syllables—allowed us to produce an astronomical number of different pieces of symbolic substance, each with its own meaning. This momentous change must have happened when our predecessors were no longer able to keep an increasingly large number of holistic vocal messages distinct from each other.

The relation between production and perception was obviously a crucial issue. Communication systems of any kind involve parity. A unit must count for the same thing for sender and receiver. It is generally accepted that the sound systems of the world's languages deal with the parity demand by means of a trade-off between a demand for perceptual distinctiveness of sounds, in order to preserve parity, and a biomechanically based (inertial) tendency toward less effort, damaging to parity, which operates in movement systems throughout the animal kingdom.

The influence of perceptual distinctiveness on the choice of production units is clearly seen in the work of Bjorn Lindblom, who was able to predict the most favored distribution in perceptual space (Formant 1-Formant 2 space) of the vowels in vowel systems with up to 9 vowels in terms of a principle of perceptual differentiation. The vowels in these systems acted like individual people positioning themselves differently in an elevator depending on where the other people were standing. Thus, the choice of particular vowels in particular languages is context-dependent (i.e., influenced by the positions of other vowels in the vowel space). This tendency, which is perceptually driven, is inconsistent with a linguistic assumption that languages choose vowels from a particular predetermined matrix in a context-independent fashion.

The role of a tendency toward articulatory ease is seen in the biomechanical limitations on early infant speech sound productions, some of the effects of which are also reflected in sound patterns of languages or in sound changes into new patterns. For example the "b" sound in *bat* and the "d" sound in *dab* are common in babbling and in languages. Also certain consonant-vowel pairings, such as tongue-front consonants like "d" with front of the tongue and vowels such as "i" as in the word *dip,* are common in infants and languages. Moreover, sound changes involving tongue-front consonants becoming newly aligned with tongue-front vowels are also common.

3. How Does It Develop?

Speechlike output begins with the babbling stage (7–12 months). This is a stage of frames without content. A typical babbled utterance is a rhythmic series of what sounds like the same consonant-vowel syllable (e.g., "bababa") without programming of separate individual content elements. In this biomechanically simple mode, the tongue tends to remain in the same position in the front-back dimension for the consonant and the vowel, as intimated earlier, and probably remains at the same height relative to the mandible.

Early in babbling, there is little or no influence on production of the detailed sound preferences of the adult language system, even though infants are known to possess relatively advanced perceptual skills at that age. Sound preferences in early babbling are the same regardless of the language environment. Beyond this time, infant sound preferences begin to converge on the specific sound preferences of the particular language they are listening to, though this tendency remains relatively slight up until the end of the first word stage (12–18 months).

4. How Does It Evolve?

One possibility is that the course of evolution of speech production is reflected in the developmental sequence. This is based on the belief that earlier hominids must have initially been biomechanically constrained in their ability to make sounds and sound sequences in much the way that modern infants are. The frame itself may have originated in the mandibular cyclicities of mammalian ingestion, (chewing, sucking, licking) and (following an intermediate communicative stage of cyclical lipsmacks, common in monkey communication) may have been paired with phonation to form protosyllabic consonant–vowel syllables.

The production–perception relation again comes to the fore when we consider speech evolution. Hominids are among a small set of vertebrates, including songbirds and some aquatic mammals, that learn their sound system. This calls for an ontogenetically malleable perception–production interface. A possible evolutionary basis for human vocal learnability is suggested by the finding, by Giacomo Rizzolatti and his colleagues, of *mirror neurons* in the monkey homolog of Broca's area and elsewhere. These are neurons that discharge when monkeys make a particular movement and when they perceive others making that movement. Though monkeys themselves do not imitate, these neurons, including perhaps the ones associated with lipsmacks, may have provided a neural substrate for the evolution of learnable speech. Using this perceptual-motor linkage capability, earlier hominids may have developed a greater capacity to evaluate the relation between communicative sounds they heard and sounds they produced, a capacity necessary for learning to produce the sounds of whichever language they were listening to.

Further reason to emphasize an evolutionarily close relation between vocal perception and production is the finding of a similar relatively direct role of the inferior (perisylvian) premotor and prefrontal cortex in humans and monkeys in the processing of vocal *input.* These regions, roughly constituting Broca's area and its monkey homolog, had previously been considered only relevant to vocal *output* in humans alone.

As to the difficult issue of the origin of the actual pairing of sound patterns with concepts, a recurring suggestion begins with the fact that about three-fourths of the words for the maternal parent in the world's languages have a nasal consonant in them. Perhaps in the presence of a nasalized infant demand vocalization directed to the maternal parent, a vocalization in common use today, the parent decided "This sound (e.g., 'mama') stands for me." The additional fact that about two-thirds of the words for the paternal parent in the world's languages do not contain a nasal consonant suggests that the first perceptual distinction between words may have used the nasal–nonnasal dichotomy.

Peter F. MacNeilage

See also Speech Perception

Further Readings

Ferrari, P. F., Gallese, P., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth movements in the monkey ventral premotor cortex. *European Journal of Neuroscience, 17,* 1703–1714.

Hardcastle, W. J., & Laver, J. (1997). *The handbook of phonetic sciences.* Oxford, UK: Blackwell.

Kent, R. D. (1997). *The speech sciences.* San Diego, CA: Singular.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge: MIT Press.

MacNeilage, P. F. (2008). *The origin of speech.* Oxford, UK: Oxford University Press.

MacNeilage, P. F., & Davis, B. L. (2000). On the origin of the internal structure of word forms. *Science, 288,* 527–531.

Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial ordering mechanism in speech production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 295–342). Hillsdale, NJ: Erlbaum.

SPEECHREADING

Speechreading, also known as lipreading, is the perception of speech by vision alone. The term *lipreading* is commonly used in nontechnical discussions and can be used interchangeably with speechreading. Speechreading is preferred when emphasizing that visual speech perception is not limited to perceiving the lips but includes perception of the motions and postures of the tongue, jaw, and cheeks, and possibly the eyebrows and head. Sometimes, speechreading refers to audiovisual speech perception by individuals with sufficient hearing loss that they rely primarily on vision but also obtain information from a hearing aid or cochlear implant. The information in this entry applies strictly to visual speech perception alone.

Visible Speech Information

The same biomechanical speech activities that produce acoustic speech information produce optical speech information. Speech production involves movements of the lips, jaw, tongue, velum (i.e., the soft palate), and glottis (i.e., the vocal folds), as well as the lungs, all of which contribute to the sounds of speech. But the speechreader has only a limited view of the speech production activities. The lips and jaw are most visible. The tongue is visible intermittently when the mouth is open. The velum, glottis, and lungs are hidden. Because all the speech production organs are controlled systematically to produce the consonants and vowels of a language, perception is impaired by partial perceptual access to their activities, as is the case for the speechreader.

For example, production of "b," "p," and "m" involves closing and then opening the lips. When these consonants occur at the beginning of an English word, they are virtually indistinguishable to the speechreader who can see the lip gestures but cannot see the additional vocal activities that generate the differences in how they sound. Specifically, "m" is a nasal consonant, because it is produced by lowering the velum and allowing air to go out through the nose, a gesture invisible to the speechreader ("n" is another nasal consonant). The distinction between "b" and "p" before a vowel is due to the timing of the lip opening relative to the onset of vibration of the glottis, which is hidden. For “b,” the vibration begins close in time to the lip opening, and for “p” the onset of vibration follows the lip opening with a perceptible delay. The relative timing of opening the vocal tract versus initiating glottal vibration is used for other consonant contrasts (e.g., “d” versus “t”) as well. Additional consonant distinctions of so-called manner are only partially visible. Manner involves the extent to which the lips or tongue block the airflow in the vocal tract, which is often difficult to see (e.g., “1” versus “d”).

Many studies have been undertaken to describe the patterns of visual perceptual confusions among consonants and among vowels. The specific confusion patterns vary somewhat depending on the talker and the skills of the speechreaders. At one extreme is the ventriloquist, whose goal is to hide all evidence of speech articulation, and at the other extreme is the oral interpreter who uses exceptionally clear visible articulation to convey another person’s speech to a deaf person. However, the consonant confusions of an average speechreader with normal hearing who views an average talker are of the following type (where the confusions are within the “{}”): {b, p, m}, {f, v}, {1, n, k, g, h}, {d, t, s, z}, {w, r}, {th (as in “thing”), dh (as in “there”}, {sh, ch, zh, j}. Vowel confusions are more difficult to list. A specialized linguistic notation is typically used, but here the confusions are listed using words to indicate exactly which vowels are mutually confused: {l*u*te, b*oo*k, b*ir*d}, {b*oa*t, b*ou*t}, {b*i*t, b*ea*t, b*e*t, b*a*ke, b*a*t}, {b*oy*}, {b*ou*ght, b*i*te, *a*bout, f*a*ther, b*u*t}.

The figure shows still images from videos of a talker saying three different vowels and three different consonants. Visible movement is critical to the speechreader, but the figure illustrates that consonant and vowel information can be seen in still images.

Figure 1 Visible Speech Gestures

Source: Communication Neuroscience Laboratory, House Ear Institute

Note: Images from videos of three different spoken vowels and two different spoken consonants.

Speechreading Proficiency

Laboratory studies have shown that speechreading can be quite accurate. The most accurate speechreaders observed to date are individuals who became deaf before acquiring language, and who rely on speechreading (not sign language) for communication. With such individuals, accuracy levels can be approximately 80 to 90% words in sentences correctly perceived. Recent research has suggested that the average prelingually deaf speechreader would score above 95% of speechreaders with normal hearing. Importantly, there is a huge range of individual differences even among deaf speechreaders: The range is from virtually zero words correct to 80 to 90% correct. The range among hearing speechreaders is from virtually zero to levels of 40 to 60% correct.

The Role of Word Patterns and Context

Given that speechreaders cannot see all of the consonant and vowel distinctions that listeners can hear, how is accurate speechreading achieved? First, skilled deaf speechreaders are able to perceive a few more vowel and consonant distinctions than the average hearing speechreader. For example, they likely distinguish between “w” and “r.” Second, the lexicon (mental dictionary) assists in limiting ambiguity for the speechreader. For example, the word *bought* exists in English, but *mought* and *pought* do not. Thus, the speechreader’s inability to distinguish among the initial consonants “b,” “p,” and “m” is not a problem for recognizing *bought*. Also, sentential semantic context can assist in disambiguating words.

Computer-based models of the English lexicon from the perspective of the speechreader have shown that many words present unique visual patterns, even after taking consonant and vowel confusions into account. One estimate showed that 68% of word patterns were still unique (not confusable with other words) after adjusting for the confusability of their constituent consonants and vowels. However, this optimistic estimate included many multisyllabic words. When only single-syllable words (e.g., *ball, hit*) were considered, the percent of unique patterns dropped to 15%. That is, most short words could be easily confused, and the number of words with which they could be confused was estimated to be on average eight. Word confusions can be resolved sometimes through context and knowledge of the conversational topic. But research suggests that the ability to perceive the consonants, vowels, and words is the primary contributor to accurate visual perception of speech.

Lynne Esther Bernstein

See also Audition: Disorders; Speech Perception; Speech Production

Further Readings

Auer, E. T., Jr. (2002). The influence of the lexicon on speech read word recognition: Contrasting segmental and lexical distinctiveness. *Psychonomic Bulletin & Review, 9*(2), 341–347.

Bernstein, L. E., Demorest, M. E., & Tucker, P. E. (2000). Speech perception without hearing. *Perception & Psychophysics, 62*(2), 233–252.

Jiang, J., Auer, E. T., Jr., Alwan, A., Keating, P. A., & Bernstein, L. E. (2007). Similarity structure in visual speech perception and optical phonetics. *Perception & Psychophysics, 69*(7), 1070–1083.

Munhall, K. G., & Vatikiotis-Bateson, E. (1998). The moving face during speech communication. In R. Campbell, B. Dodd, & E. Burnham (Eds.), *Hearing by eye: Advances in the psychology of speechreading and auditory-visual speech* (Vol. 2, pp. 123–142). Hove, UK: Psychology Press.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26*(2), 212–215.

SPEED OF PROCESSING IN SENSORY SYSTEMS

Psychologist Irving Biederman suggested the following experiment: Turn the television on without the sound, change channels with your eyes closed and then briefly open your eyes. In all likelihood, your brain can effortlessly analyze the image on the screen, even if you have no idea what will be there. It even seems as though the processing is instantaneous, but clearly this cannot be the case. It certainly must take some time for the information to be processed by the visual system. But how can researchers measure how long it takes to do this processing? And what can such information tell us about the underlying mechanisms? These questions are the focus of this entry.

Measurement and Underlying Processes

The most obvious approach is to use some measurement of behavioral reaction time. This might involve determining how long it takes to name an object by using a voice key to record the moment at which the subject starts to vocalize. However, this type of responding takes quite a long time—at least 600 milliseconds (ms)—probably because accessing the verbal system is time consuming. There is also a large effect of word frequency on naming time—it is faster to name an "apple" than a "xylophone" for example—but it seems unlikely that this reflects a difference in sensory processing time.

An alternative approach would be to use a manual response, such as key pressing. For example, if a subject is asked to press a button each time a picture of an animal is flashed on a screen, minimal reaction times can be as short as 280 to 300 ms. But such values include not just sensory processing, but also the time needed to initiate and execute the motor response. How can researchers get a more direct measure of sensory processing speed?

One possibility is to look for brain signals in the electroencephalogram (EEG) or using magnetoencephalography (MEG) related to processing, but that precedes any behavioral responses. For example, in the animal detection task, there is a clear and robust difference in the event-related potentials (ERPs) generated on target and distractor trials starting about 150 ms after stimulus onset, implying that the underlying processing has been done by then. Although visible at frontal recording sites, it is now clear that the signals are actually generated in more posterior visual areas. Importantly, this differential response occurs even for images that the subject has never seen before and despite the fact that there is no way to predict what sort of animal will be shown.

Is this value of 150 ms a reasonable estimate of the speed of visual recognition? Intriguingly, recent experiments using eye movements instead of a hand response suggest that processing may be even faster. When two images are simultaneously flashed left and right of fixation, subjects can make reliable saccades (rapid eye movements) to the side where there is an animal in as little as 120 to 130 ms—substantially before the 150 ms differential EEG effect. If we allow at least 20 ms to initiate the eye movement, this implies that the presence of an animal in an image can be detected in no more than 100 ms. Such numbers put severe constraints on the underlying mechanisms.

Note that there may be important differences between these extremely rapid eye movement responses and more classic behavioral responses measured with key presses. One feature of key-pressing tasks is that subjects can easily modify the target category from block to block, and even trial to trial. Thus, in an experiment where pictures of animal faces and human faces are mixed, subjects can easily respond just on animal trials, but then switch to responding only to humans in the next block of trials—scoring over 97% correct for both types of target. In contrast, the very fast saccadic eye movement responses to animals and faces appear to be much harder to control. The fact that our eyes appear to naturally move toward stimuli that have a high salience (including faces and animals) seems sensible because in one sense, there is no such thing as a “wrong” eye movement (except in artificial laboratory situations where you might get punished with the sound of a buzzer for getting it wrong!). However, it would be a bad idea to make a manual response without first checking that you have made the right decision. These differences could explain why manual choice reaction times are so much longer.

Although behavioral measures such as reaction times have provided much useful information about processing speed, an even more direct method is to record the activity of single neurons in the brain and measure when they start to respond following the presentation of a stimulus. For example, in the monkey visual system, the earliest spiking responses at the output of the retina occur with latencies of about 20 ms, and these latencies increase progressively throughout the visual pathways, reaching 80 to 100 ms in areas such as the inferotemporal cortex (IT). This area lies at the end of the so-called ventral processing stream and contains neurons with highly selective responses to key stimuli, such as faces.

The speed of these responses has important implications for the nature of the underlying processing. For example, if we count the number of layers of neurons between the photoreceptors in the retina and face-selective cells in monkey IT, it appears that visual information needs to cross around 10 sets of synapses in only 100 ms. This leaves only about 10 ms for processing at each stage before the next layer needs to respond; a fact that severely constrains the amount of information processing that can be performed and limiting the possibilities for using feedback loops. Significantly, the neurons can be selective right from the exact moment they start to respond, a feature that is one of the hallmarks of feedforward processing. Another observation consistent with feedforward processing is the fact that IT neurons can still respond selectively to rapidly changing images, as in rapid serial visual presentation (RSVP) paradigms. The finding that at least some aspects of high-level vision can be done using only feedforward mechanisms is a surprising one. It is well known that there are feedback connections at virtually every stage of the visual system (the only exception being the retina, which does not receive feedback from the first relay station in the thalamus), and it is often assumed that these feedback connections are required for all types of visual processing.

Factors Influencing Response Latencies

What factors influence the latencies at which visual neurons respond? An obvious one is the intensity of the visual stimulus itself, because simply lowering contrast can produce very large increases in response latency. This finding is easy to understand when you consider the “integrate-and-fire” nature of neurons—it takes longer for a neuron to charge up and reach a threshold when the stimulus is weak. This fits with behavioral measures of reaction time that also increase in a characteristic way with decreasing stimulus energy. But the nature of the stimulus is also important in determining the latency of the neuronal response. If the stimulus is not well matched with the receptive field characteristics of the neuron, the latency of the neuron will also be increased and the neuron may not even respond at all.

Familiarity may also be an important factor in determining latency. In one recent study, it was found that face-selective neurons in monkey IT responded about 20 ms faster to faces of humans and monkeys than to the less familiar faces of other animals, such as birds and cats. Importantly, this was not simply the result of the neurons having a weaker response to those other stimuli. Could it be that the particularly rapid responses to monkey and human faces reflect the fact the monkeys have had far more experience with those stimuli? This is an interesting question that deserves further study.

While most single unit recording studies have been performed in animals, there have been a number of fascinating recent studies of single neuron activity in human patients undergoing surgery for intractable epilepsy. Neurons in temporal lobe structures such as the hippocampus have been reported that respond selectively to particular individuals or places well known to the patient ("Bill Clinton," "Halle Berry," or "The Taj Mahal"). Remarkably, such neurons can respond to effectively any image that corresponds to the object, including, for example, the name "Halle Berry" written as text! Clearly, the existence of such neurons presumably reflects the extensive experience that the patients have with those particular stimuli, and indeed, these highly selective responses have only been seen for stimuli that are well known to the subject. From the point of view of processing speed, it is important to note that the onset latencies of such neurons are often relatively long (280–300 ms), a result compatible with the idea that the responses are perhaps related more to mnemonic functions than sensory processing per se. So far, there is relatively little data on response latencies in the cortical areas that provide the visual inputs to these neurons, and it will be interesting to see whether highly selective responses can also be found there.

When comparing processing speeds and response latencies in humans and other animals, it is important to realize that humans are systematically slower than monkeys on virtually whatever task you try them on. This has been termed the two-thirds rule—monkeys take two-thirds of the time taken by humans in the same task. For example, rhesus monkeys can respond to the presentation of an animal in as little as 160 to 180 ms, substantially shorter than the minimum latency for humans. It might be thought that this could be because visual processing in humans is somehow more sophisticated. But in fact, monkeys show remarkably similar patterns of responding to humans. Instead, it seems likely that the difference is essentially related to brain size. A lot of the delays in the input–output pathway are probably related to the conduction velocities of intracortical connections. Estimates put the value at around 1 to 2 meters per second, which means that a substantial part of the difference in latency between V1 and IT in the monkey may simply be the result of conduction delays—the distance is about 30 mm in monkeys, which could easily result in a delay of up to 30 ms.

Other Sensory Systems

So far, this entry has been discussing primarily processing speed in the visual system. But, of course, similar questions can be asked for other sensory systems. In some ways, vision is relatively slow, simply because of the fact that transduction in the photoreceptors of the retina is fairly slow. Virtually no spikes leave the retina in less than about 20 ms. In contrast, responses in the auditory and somatosensory systems can be extremely fast, with the first spikes leaving the sensory receptors at latencies of just a couple of milliseconds. Indeed, the responses of neurons in bat auditory cortex, some 6 synapses away from the receptor cells, can be as little as 12 ms, leaving only 2 ms for processing at each stage.

All of these temporal constraints put severe limits on the sorts of coding that can be used in the nervous system. In particular, it makes it difficult to imagine that the nervous system would have time to get an accurate measure of the firing rate of individual neurons. Instead, it seems likely that the brain makes use of the relative timing of spikes across entire populations of neurons to achieve the speed of processing needed for survival. The fact that the auditory system uses relative timing of spikes is well known, and indeed our ability to localize sounds depends at least in part on the fact that sounds can arrive with a delay of around 0.5 ms between the two ears when the sound is presented off to one side. Similar processing mechanisms that make use of small timing differences could also apply in other sensory pathways.

In general, it is clear that the speed with which the brain can process stimuli is absolutely critical for survival, whether you are predator or prey. It is remarkable that the brain, with neurons that fire, at most, a few hundred times per second and with conduction velocities for nervous transmission that are so slow, can compete with and even outperform the most sophisticated artificial vision systems.

Simon J. Thorpe

See also Consciousness; Neural Representation/Coding; Priming; Rapid Serial Visual Presentation

Further Readings

Hegde, J. (2008). Time course of visual perception: Coarse-to-fine processing and beyond. *Progress in Neurobiology, 84*(4), 405–439.

Kiani, R., Esteky, H., & Tanaka, K. (2005). Differences in onset latency of macaque inferotemporal neural responses to primate and non-primate faces. *Journal of Neurophysiology, 94*(2), 1587–1596.

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research, 46*(11), 1762–1776.

Mormann, F., Kornblith, S., Quiroga, R. Q., Kraskov, A., Cerf, M., Fried, I., et al. (2008). Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *Journal of Neuroscience, 28*(36), 8865–8872.

Quian Quiroga, R., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature, 435,* 1102–1107.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*(6582), 520–522.

STATISTICAL LEARNING

It is natural to think of perception in terms of the processing of individual features (such as color and shape) and how they are combined into discrete objects (such as animals and bicycles). This simple characterization underestimates the information that is available in perceptual input, though, because there are also massive amounts of information about how these features and objects are *distributed* in space and time. In time, for example, eating food at a restaurant is more likely to be followed by paying a bill than by climbing a tree—just as (in English) the syllable /sci/ is more likely to be followed by /ence/ than by /on/. And in space, for example, a car is more likely to be next to a bicycle than to a stapler. Discovering such regularities is difficult because they are embedded within complex and continuous environments where not all information is relevant. But the mind is nevertheless sensitive to such regularities, uncovering them in part by means of *statistical learning:* an automatic and unconscious perceptual process that encodes statistical regularities across space and time. We are often unaware of the operation of statistical learning, yet it may play a crucial role in segmenting the continuous perceptual world into discrete manageable units, such as words, events, objects, and scenes. This entry describes statistical learning, how and when it operates, and how it may support online perception.

Segmenting the World via Statistical Regularities

Statistical learning (henceforth abbreviated SL) is a type of *implicit* learning, in that it can occur without intent or awareness. Indeed, like many other aspects of perception, it occurs beneath the level of consciousness, and we are not even normally aware that our minds are engaging in such learning at all. Implicit learning has a long history in psychology, but the study of SL in perception research has arisen only in the last 15 years, as researchers have focused on how such processing may serve to segment continuous perceptual input into discrete units.

Auditory Statistical Learning

In its modern incarnation, the study of SL began in the domain of language acquisition. When you first hear a foreign language, you do not understand the words, but you also may find it difficult to tell where the words start and stop in the first place. This is because most natural speech involves a largely *continuous* pattern of sound: What seem like pauses between individual words do not exist in the sounds entering our ears, but are rather constructed by our minds. This means that a first challenge for language learners is simply to find word boundaries in natural speech. Developmental researchers including Richard Aslin, Elissa Newport, and Jenny Saffran proposed that this task might be accomplished in part by the processing of statistical regularities in speech streams. Using a simple artificial language, they showed that eight-month-old infants were able to use statistical regularities over time to segment words out of continuous syllable streams.

Infants in such experiments might initially hear a continuous stream of syllables consisting of repeated sequences of four different triplet "words," each composed of three syllables (e.g., *bidaku, golabu, padoti, turopi*). These four triplets are repeated in a randomized order in a continuous stream, with no acoustic cues (such as longer pauses or prosody) to indicate where each triplet starts or stops (e.g. " . . . *bidakupadotigolabubidakuturopi* . . ."). Nevertheless, this stream contains robust statistical structure: The first syllable in a triplet perfectly predicts the second syllable, but the third syllable could be followed by several other syllables (corresponding to the first syllable of the following word). For example, if we denote the four triplets as *ABC, DEF, GHI,* and *JKL*—with each syllable denoted by a capital letter—then *A* will always be followed by *B,* but *C* will be followed by either *D, G,* or *J.* After only a few minutes of exposure to such a stream, a test phase then begins in which the infants are presented with three-syllable sequences in isolation—either triplets (e.g. *ABC*) or "nonword" foils composed from the same syllables, but in an order they have never heard before (e.g. *AEI*). Despite the fact that all of the individual syllables are equally familiar, infants can reliably discriminate the words and nonwords (as evaluated by a procedure that measures how long they attended to each type). Moreover, they can even discriminate triplets (e.g. *ABC*) from "partword" foils that they did hear, but less often (e.g. *BCG,* from when triplet *ABC* happened to be followed by *GHI*)—and later studies demonstrated a sensitivity to even more subtle statistical patterns. This is an example of SL because (1) the relevant regularities existed only in the distribution of syllables in time, because each individual syllable was heard equally often, and (2) the triplet "words" were obscured in an otherwise continuous stream of perceptual input.

This ability has led some researchers to speculate that SL may be a means by which children come to learn where the words are in continuous speech, but the precise relationship of SL to language acquisition remains uncertain. On one hand, this connection is supported by demonstrations that SL occurs more readily for linguistically relevant auditory information, and that it directly facilitates subsequent word learning. On the other hand, SL operates in subjects who have already mastered language (i.e., human adults), and even in nonlinguistic subjects (including nonhuman primates and even rodents)—and it also operates over many types of nonlinguistic input, such as musical tones. Moreover, some computational modeling studies suggest that SL is not sufficient to identify words in actual natural-language speech streams.

Visual Statistical Learning

Though SL had its origins in studies of language acquisition, it also operates pervasively in visual perception, where it has recently become a phenomenon of considerable interest. Visual statistical learning (VSL) can be measured in adult observers, for example, with continuous sequences of nonsense-shapes, when they appear one after another in an exact analogue to the auditory studies with infants—with each auditory syllable now replaced by a visual shape. After passively observing such a shape sequence (always with a constant between-shape delay) for several minutes, observers are able to reliably distinguish triplet "words" from three-shape foils when they are pitted directly against each other (now explicitly segmented) in a surprise "Which is more familiar?" test. Recent functional magnetic resonance imaging (fMRI) studies using similar designs have been able to explore the time course of VSL, and suggest that it is highly efficient—with the brain already discovering the presence of statistical structure after only the third or fourth repetition of each triplet.

VSL can also extract spatial regularities when multiple objects are viewed simultaneously. In some such studies, for example, observers see several shapes that are placed into a grid in reliable patterns (unbeknownst to them). For example, shape A might always appear immediately above and to the right of shape B. Because many such patterns are spatially interleaved in each grid, however, the pairs can only be spatially segmented on the basis of their statistically reliable spatial relations as observers see grid after grid. At test, observers are presented with the original spatial pairs pitted against foils consisting of two equally familiar shapes that had appeared together less frequently. Again, observers are able to judge the actual pairs as being more familiar.

Note that although SL may operate over spatial or temporal regularities, some of the mental representations that result from SL may be abstracted from such contexts. Indeed, VSL of spatial layouts can later be expressed in purely temporal contexts, and VSL of temporal sequences can later be expressed in purely spatial layouts.

What Type of Process Is Statistical Learning?

Since SL was first studied in these ways in the mid-1990s, a great deal has been learned about its underlying nature.

Automaticity and Implicitness

In adults, SL has most commonly been measured by explicitly separating sequences into their respective triplet "words" in a separate test phase, and then testing to see whether observers judge them to be more familiar than various types of foils, in forced choices. Such familiarity judgments could indicate conscious recognition of the statistical structure, but in fact such judgments are often reliable even when subjects were completing a separate "cover" task during the initial exposure, and even when they think they are merely guessing during the test phase. VSL has also been demonstrated with several implicit measures, however, including behavioral measures showing speeded response times for statistically predictable targets, and fMRI measures that have shown robust neural sensitivity to statistical structure even in subjects who show no SL by conventional familiarity tests. At the same time, additional studies have shown that SL in both the auditory and visual domains will only operate when the relevant stimuli are attended. If attention is instead directed away to other stimuli during the initial exposure, SL will not occur. Thus, it seems that SL is automatic in some senses, but not in others: It is gated by attention, but it nevertheless operates without intent or awareness.

The Input to Statistical Learning

One of the most critical steps in understanding any perceptual process is to determine the types of input over which that process operates. SL operates in multiple sensory modalities (including touch, in addition to vision and audition), and can accommodate many kinds of individual stimuli (such as musical tones in audition and dynamic events in vision). Research on VSL of temporal sequences also indicates that learning can operate at almost every level of the "hierarchy" of visual processing—from low-level visual features (such as shape), to discrete objects containing multiple features (such as color and shape), to high-level categories of visual scenes (such as "kitchens" and "forests").

Moreover, SL may help to determine what counts as an "object" in the first place. Statistical associations over time, for example, seem to help determine that a visual feature, such as color, is an intrinsic part of some objects (e.g., bananas) but not others (e.g., t-shirts). As a result, SL for colored-shape sequences can be expressed later for monochromatic shapes if the color-shape associations were loose during the initial exposure, but not if color and shape were reliably paired. For example, if shapes are assigned to unique fixed colors, then learning of colored-shape triplets can be expressed when the shapes are presented with their original colors at test, but not when the same shapes are presented in black during test (a context that is sufficient to express learning when the shapes have colors that randomly vary during familiarization). Similarly, studies of spatial VSL have demonstrated that learning does not operate efficiently over spatial patterns that are embedded within larger reliable patterns. For example, when spatial triplets are embedded in static grids, observers learn the triplets, but not their component pairs. And in audition, learning of *nonadjacent* statistical dependencies in temporal sequences is usually difficult, but it becomes easier when the interrupting stimuli are either highly variable or perfectly constant (as opposed to when they are only moderately variable). In each case, statistical variability seems to determine the "chunks" over which learning operates in the first place.

Controlling Statistical Learning

Recent research has also begun to determine the conditions under which SL will cease to operate. In most temporal SL studies, for example, the regularities are present in the sequences as soon as they begin, and they persist until the initial exposure phase ends. Learning is influenced in interesting ways, however, when the regularities appear only at the beginning or ending portions of the sequences—with the rest of the sequences filled with completely random orders of syllables or shapes. (Because of the implicit nature of SL, subjects in such experiments are not overtly aware of such transitions.) In this situation, learning will still occur when the structure precedes the structure-less "noise": the regularities that are initially discovered may remain robustly encoded, even when they later "fall apart." However, when the noise precedes the structure, learning is either weak or nonexistent: Perceptual processes may effectively learn that there is "nothing to learn" based on the initial noise and will stop trying (even though subjects still attend to the sequence).

What Is Statistical Learning Good For?

Statistical learning appears to a ubiquitous process in perception, but what is it good for? Recent research suggests at least three answers. First, by segmenting continuous streams of input, SL may effectively form the "chunks" that later processes depend on (e.g., the "words" in syntactic processing or the "objects" stored in visual working memory). This is supported by demonstrations that SL of linguistic stimuli can directly facilitate later word learning or relative-clause comprehension. Second, SL may speed later perceptual processing, making it more efficient—as is apparent in implicit response-time measures of VSL. Third, SL may be adaptive in that it yields a type of "future-oriented" processing, helping us to predict what we may be about to experience, so that we can adaptively tune our current behavior. Recent fMRI studies of VSL, for example, suggest that it yields a type of "implicit anticipation" of upcoming stimuli. Such implicit anticipation may be responsible for the facilitation previously described and may help the visual system cope with ambiguous or degraded stimuli. In sum, statistical learning is clearly a product of perception, but it may also in turn facilitate perception.

Brian J. Scholl and Nicholas B. Turk-Browne

See also Perceptual Learning; Unconscious Processes; Visual Scene Statistics

Further Readings

Fiser, J. Z., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12,* 499–504.

Fiser, J. Z., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28,* 458–467.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10,* 233–238.

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science, 12,* 110–114.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926–1928.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134,* 552–564.

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (in press). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience.*

Yang, C. D. (2004). Universal grammar, statistics, or both? *Trends in Cognitive Sciences, 8,* 451–456.

SURFACE AND MATERIAL PROPERTIES PERCEPTION

Objects in the environment can be described at many scales. For example, an orange has a shape (roughly round) and, if we look more closely, we notice that the surface is smooth but irregular, composed of small bumps. If we continue to a microscopic scale, the interaction of light with the molecular structure of the surface of the orange gives it a characteristic shiny or glossy appearance and a characteristic color. From a physicist's point of view, the surface exists at many different spatial scales: megascale, mesoscale, and microscale, according to a classification adopted by Jan Koenderink and Andrea van Doorn.

The visual system provides considerable information about surfaces at different scales (small, round, rough, glossy, orange, etc.) that correspond roughly—but only roughly—to identifiable physical properties of the surface. The perceptual properties that correspond to megascale physical properties signal information about the size, shape, and orientation of objects. This entry is primarily concerned with perceptual properties such as roughness or glossiness that correspond to mesoscale and microscale properties of objects other than their color. These perceptual properties provide information about the material that the object is composed of; the differences that signal expensive ceramic or tacky plastic or that mark the safe places to walk on an icy day.

The glossiness or color of a surface is only available visually. Its roughness can be estimated either visually or haptically (through the sense of touch) and the correspondence between haptic and visual properties is potentially complex. Haptic properties, like visual properties, exist at multiple spatial scales. A knitted sweater serves as an example where both haptically and visually we can separately estimate the ridged pattern of yarn and the fuzziness of the fibers of which they are composed.

The two central research issues concerning surface material perception are, first of all, to characterize perceptual properties such as glossiness in physical terms (the psychophysical problem), and second, to assess how reliably the visual system can assess any particular material property (the constancy problem).

The Psychophysical Problem: Glossiness

Suppose that we place a large number of small objects (shoes, ships, sealing wax) on a table and then ask an observer to order them according to their glossiness (or shininess). The surfaces can also vary in roughness (mesoscale), color (microscale), or any other surface material properties but the surface material of each should be homogeneous.

Let's imagine that the observer eventually arrives at an ordering by glossiness of all of the materials on the table that he judges satisfactory. The question for us now is to characterize physical measures that correctly account for his ordering and predict where he might insert a new object in the sequence. For convenience, we will use the term "gloss" to refer to the (unknown) physical property corresponding to perceived "glossiness" in discussing various conjectures as to what "gloss" might be.

There are industry-standard physical definitions of surface gloss based on analyses of how surfaces reflect light arriving from different directions. Researchers have measured the psychophysical function relating to the perceived glossiness of surface samples that are uniformly spaced along an industry-standard gloss scale. Gaël Obein and colleagues, for example, used a novel scaling method to measure this psychophysical function (relating glossiness to gloss) for such a series of test surfaces. They measured it under different lighting conditions to see how perceived glossiness varied with lighting condition and they had observers judge the glossiness of surfaces viewed with one eye only and with both. In testing whether perception of glossiness is invariant under changes in lighting conditions, they are testing whether the visual system has some degree of glossiness constancy analogous to color constancy.

They found that constancy failed: Changes in lighting and viewing conditions affected perceived glossiness. However, the shape of the psychophysical function remained roughly the same across conditions. Observers were most sensitive to differences at the two extremes of the physical scale (nonglossy and mirrorlike) and less sensitive in between.

The research of Obein and colleagues does not tell us exactly what physical measure determines perceived glossiness for surfaces other than the industry-standard reference surfaces that they used. Isamu Motoyoshi and colleagues proposed such a measure based on image statistics, and the most remarkable aspect of their proposal is its simplicity. They created a series of stuccolike surfaces whose surface material properties were systematically varied. The surfaces were coated uniformly with both a matte (nonglossy) paint and a layer of (glossy) acrylic media coating. Surfaces differed in **albedo** (the proportion of light reflected) of the matte paint or density of the acrylic coat. Intuitively, variations in the albedo of the matte paint were intended to influence perceived lightness of the surface, whereas the density of the acrylic media coating was expected to affect its perceived glossiness. The surfaces shared the same stuccolike spatial structure.

Motoyoshi and colleagues first analyzed the histograms of intensity values of black and white photographs of the surfaces. If we normalize the intensities of pixels so that they fall in the range 0 to 255 rounded to integer values, then the histogram corresponds to the count of pixels with intensity 0, with intensity 1, and so on, independent of spatial structure. This sort of histogram is commonly found in programs for manipulating images on computers. Motoyoshi and colleagues investigated the relationship between simple statistics of the histogram and the perceptual properties of the surface. For example, we might expect that the standard deviation of the histogram of an image to correspond roughly to its perceived contrast.

Motoyoshi and colleagues analyzed the histograms of the images of their stimuli and conjectured that an image statistic called *skewness* was a good predictor of how glossy a surface appeared. Skewness is a measure of deviation from a symmetric, bell-shaped histogram with positive skewness corresponding to a histogram that tails off slowly to the left but drops abruptly on the right, and negative skewness describing the opposite pattern. Motoyoshi and colleagues asked observers to rate the glossiness of each of their stimuli and found that mean rating of glossiness was close to a linear transformation of skewness. Intuitively, histograms with high skewness typically, but not always, correspond to images with a concentration of bright pixels.

They also transformed images of stucco, black cotton fabric, and crumpled white paper, varying the skewness of the histograms of each image. Observers rating the glossiness of the transformed images increased or decreased as skewness was artificially increased or decreased. Motoyoshi and colleagues also showed that adaptation to images that have skewed histograms leads to changes in perceived glossiness of subsequently viewed histograms.

The Constancy Problem: Interactions

Motoyoshi and colleagues studied roughness or shininess with other surface properties held constant. In our thought experiment, we asked the observer to compare the glossiness of objects (shoes, ships) on the experimental table that could differ in color or roughness as well as glossiness. Similarly, we could ask observers to judge the roughness or color of objects that differ in glossiness or other surface material properties. If observers' judgments of glossiness are little affected by changes in roughness, for example, then observers exhibit a perceptual *constancy*. Intuitively, the "constant" visual system estimates glossiness without "contamination" from surface properties that vary with "roughness."

Yun-Xian Ho and colleagues used a method called *additive conjoint measurement* to derive scales for the perceptual correlates of two surface-material properties, physical gloss and physical roughness (which they refer to as "bumpiness") and to test whether physical glossiness affected perceived roughness or physical roughness affected perceived glossiness. Their stimuli were computer-generated, binocularly presented images of irregular surfaces composed of many superimposed "bumps," and the average height of the bumps corresponded to physical roughness. Physical glossiness was manipulated by variation in a computer-graphic graphics parameter that affected the coherence of reflected light. Their stimuli consisted of a five by five array of surfaces whose rows differed in physical roughness and whose columns differed in physical glossiness. Rather than simply test whether each physical roughness affected perceived glossiness or vice versa, they estimated the interaction between the two properties and modeled it.

In one condition ("roughness"), observers were asked to compare pairs of surfaces and judge which was rougher. In the other condition ("glossiness"), different observers saw the same pairs but judged which was glossier. Ho and colleagues used additive conjoint measurement to measure the contributions of physical roughness and physical gloss to perceived roughness in the roughness condition. Similarly, they measured the contributions of physical roughness and physical gloss to perceived glossiness in the glossiness condition. Of course, for the "constant" observer, variations in gloss would not affect perceived roughness, and variations in physical roughness would not affect perceived gloss. The observers in both conditions were, in effect, asked to "tune" their perception to either mesoscale (roughness) or microscale (gloss) and ignore physical properties at other scales, much as a good radio can be tuned to receive one station while excluding others.

Ho and colleagues found that most human observers perceived physically glossier surfaces to be rougher and physically rougher surfaces to be glossier. However, when judging roughness, the effect of changes in physical roughness were markedly attenuated and vice versa. The difference in perceived roughness varied about 10% as gloss was varied across its full range, and the difference in perceived glossiness varied by about 30% across the range of roughness considered. Observers exhibited partial but not perfect constancy.

The Constancy Problem: Illumination

The illumination in any scene can vary, and visual judgments of material properties are potentially affected by changes in illumination. Such effects are examples of failures of a second constancy; a constancy with respect to illumination, already mentioned in describing the work of Obein and colleagues. In a second study, Ho and colleagues examined how the direction of illumination affected the perceived roughness of computer-generated, binocularly presented images of irregular surfaces. The surfaces resembled pieces of sandpaper and differed in the average height of the irregularities (defined to be physical roughness) in the surfaces. The experimenters varied both the physical roughness of surfaces and the direction to the light source in the computer-generated scenes. Observers could readily judge relative roughness under any single illumination direction condition, but changes in illuminant direction led to systematic changes in perceived roughness. As the direction of illumination deviated from perpendicular to the surface, the perceived roughness of the surface increased. If we imagine the light source as the sun and the surface placed horizontally on the ground, the same surface was judged rougher as the sun approached sunset than when it was overhead at noon.

Ho and colleagues interpreted their result as a failure of cue combination. There are potentially many sources of information ("cues") that might signal information about the roughness of a surface. Some of these, such as the proportion of the scene in shadow, are visual and vary with light position as well as with surface roughness. When the angle between surface and light direction is small (and the light direction is closer to parallel with respect to the surfaces) shadows grow in length even though the physical roughness of the surface has not changed. However, if the irregular elements composing the surface increase in height with the light direction not perpendicular to the surface, then shadows grow as well and the increase in shadow corresponds to an increase in roughness. Ho and colleagues conjecture that the visual system has mistakenly made use of cues such as proportion of shadowed area that, in the context of their experiment, are misleading indicators of physical surface roughness. They have caught the visual system doing something clever but wrong.

Ho and colleagues conjectured that such "pseudo-cues" may arise in the context of everyday experience, and it is possible that further investigation of visual and haptic cues to surface material properties will disclose further examples of pseudo-cues. For example, glossier objects may be judged harder than similar objects that are not glossy simply because glossy objects have tended to be harder than nonglossy objects in our experience. Such failures in perception of surface material properties could reveal much about how the visuomotor system combines haptic and visual information to arrive at estimates of objective properties of the environment.

Laurence T. Maloney

See also Haptics; Temperature Perception; Texture Perception: Tactile; Texture Perception: Visual; Vibratory Perception; Vision

Further Readings

Ho, Y.-X., Landy, M. S., & Maloney, L. T. (2006). How illuminant direction affects perceived visual roughness, *Journal of Vision, 6,* 634–648.

Ho, Y.-H., Landy, M. S., & Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture. *Psychological Science, 19*(2), 196–204.

Koenderink, J. J., & Van Doorn, A. (1996). Illuminance texture due to surface mesostructure. *Journal of the Optical Society of America A, 13,* 452–463.

Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature, 447,* 206–209.

Obein, G., Knoblauch, K., & Viénot, F. (2004). Difference scaling of gloss: Nonlinearity, binocularity, and constancy. *Journal of Vision, 4,* 711–720.

SYNESTHESIA

For people with *synesthesia,* ordinary stimuli such as music, numbers, or words are imbued with an "extra" perceptual quality not shared by most members of the population. Music may trigger swirling shapes, numbers may have color, or words may evoke localized taste sensations felt in the mouth. For people who have synesthesia, such experiences are considered completely normal, and they can't imagine a world without them. The absence of any obvious outward manifestation of having synesthesia has contributed to its historical obscurity in scientific research. However, much of the contemporary research is interested in the cognitive and neural basis of conscious perceptual experiences, and synesthesia is back on the scientific agenda. Synesthesia also turns out to be far more common than previously believed. Recent prevalence estimates suggest that colored letters and numbers (grapheme-color synesthesia) affects 1 to 2% of the population, with some form of synesthesia affecting 1 of the 5 classical senses in 4.4% of the population. This high prevalence has led people to speculate on the evolutionary significance of synesthesia. This entry describes the key characteristics and varieties of synesthesia, neural and genetic basis of synesthesia, and some reasons why synesthesia exists.

Key Characteristics

Synesthesia has a number of core characteristics that, collectively taken, constitute a more formal definition of synesthesia that is agreed upon by most researchers in the field. First, the experiences are normally defined as being perceptual in nature. For example, a grapheme-color synesthete *sees* that a letter A evokes red rather than remembers "red" as a verbal memory label. Functional imaging studies are consistent with this view. However, there are some borderline examples. Not all synesthetes see their colors externally, but rather in their mind's eye. Some researchers also extend the definition of synesthesia to nonperceptual associations, such as attributing genders and personalities to graphemes (e.g., "3 is male and a jerk"). This could also be construed as a related characteristic (e.g., due to more widespread changes in the brains of synesthetes), rather than a type of synesthesia in its own right. Second, synesthesia is elicited. The synesthetic experience (known as the concurrent) is elicited by some stimulus (known as the inducer). This distinguishes synesthesia from hallucinations, in which there is typically no known inducer. As such, all types of synesthesia can be named in terms of an inducer–concurrent pair (e.g., grapheme–color, music–vision, lexical–gustatory). It is to be noted that the inducer is not always an externally perceived stimulus. Thus, thinking of an arithmetical sum (e.g., 5 + 2) can trigger a synesthetic color, and thinking of an unretrievable word in a tip-of-the-tongue state can trigger a taste. The relationship between the inducer and the concurrent is normally stable over time, and it is normally unidirectional. Seeing a letter 'A' may trigger red, but seeing red doesn't elicit a percept of the letter A (however, there is more convincing evidence for *implicit* bidirectionality in synesthesia). Finally, synesthesia is assumed to occur automatically. It cannot be switched on or off at will, and this distinguishes synesthesia from most forms of imagery. Experimentally, this has been extensively studied using a synesthetic variant of the Stroop test. In the original version of the Stroop test, people are slower at naming the ink color of a word (e.g., RED) if it is incongruently colored (e.g., in green ink) than congruently colored (i.e., red ink). In the synesthetic version of the Stroop test, synesthetes show a similar effect if the grapheme is printed in an incongruent color to their synesthesia. Interestingly, there appears to be greater interference if opponent colors are used (e.g., a synesthetically red grapheme printed in green) than non-opponent color (e.g., a synesthetically red grapheme printed in yellow), suggesting that the interference has a perceptual rather than purely mnemonic component to it.

The Varieties of Synesthesia

Although there are some common principles that are common to all (or almost all) types of synesthesia, there is a large range of different combinations of inducer–concurrent pairings. These are now briefly summarized.

Grapheme–Color Synesthesia

This variety of synesthesia has been extensively studied. Top-down influences play a large role in determining the color, such that ambiguous graphemes take on the color implied by their context (e.g., a vertical line may take on either the color of the letter I or the number 1, depending on whether it is surrounded by letters or numbers). Bottom-up influences (e.g., particular fonts) may also affect the intensity/saturation of the hue. When learning a second alphabet (e.g., Cyrillic), the colors normally migrate across based on visual and/or phonetic similarity.

Some synesthetes see the color "out there on the page" like a colored overlay (these have been termed *projectors*), whereas others see it elsewhere (termed *associators*), often in their mind's eye. One suggestion is that projected colors are elicited earlier in the visual stream. A similar suggestion is that for some grapheme–color synesthetes, the color is elicited by "lower" levels of processing (e.g., the grapheme shape and early color processing), whereas for others it is elicited by "higher" levels (e.g., the meaning of numbers and possibly later stages of color processing).

Colored-Hearing Synesthesia

The term *colored hearing* was historically used as a synonym for synesthesia (*farbenhören* in German, *audition colorée* in French). In reality, the term denotes a number of distinct subtypes such as music–vision (or indeed nonmusical sounds) and speech–color, in which spoken language (but not other heard sounds) trigger color. With regard to the former, the visual experiences from music consist not only of color, but also of movement, shape, size, and location. With regard to the latter, it is typically the first *letter* of a spoken word that determines its color. Thus, *photo* and *people* would tend to be colored the same (because of the letter P) but *photo* and *fish* would not (despite the shared initial phoneme). These synesthetes often perceive the spoken word spelled out like ticker tape. This suggests a two-stage process by which phonemes are converted to letter strings and the letters trigger the color. The apparent rarity of true phoneme–color synesthesia has implications for theories of synesthesia.

Synesthesia for Spatial Sequences

Ordered sequences such as days, months, numbers, and the alphabet are often perceived as arranged in a particular sequence in space. Numbers are typically arranged from left-to-right, at least in European and North American synesthetes, although they can also twist and turn through any direction of space. This, to some extent, resembles the normal spatial bias in numerical processing found in people who lack synesthesia.

Synesthesia Involving Taste and Smell

These forms of synesthesia are rarer. Richard Cytowic documented the case of MW, "the man who tasted shapes," for whom flavors elicited tactile shapes that could be felt and explored on the hands. Synesthesia in which taste or flavors are the concurrent experience have been noted to be triggered by words that are either read or heard (e.g., *New York* tasting of runny egg yolk). In all cases in which words trigger tastes, it has been observed that similar sounding words tend to elicit similar tastes, suggesting that there is an underlying logic to the associations.

Synesthesia Involving Touch and Pain

Pain has been noted to induce synesthetic visual experiences. Synesthesia in which tactile experiences are concurrent were, until recently, considered to be rare, although recent research casts doubt on this. In particular, a novel variety, so-called mirror-touch synesthesia, has been documented in which observing touch to another human (but not object) triggers somatotopically appropriate sensations on the perceiver's own body. This has been demonstrated both with functional magnetic resonance imaging (fMRI) and using an interference paradigm in which felt touch must be discriminated from observed touch.

Neural and Genetic Basis

Functional Imaging and Electrophysiological Studies

A number of functional imaging studies have shown activity in human area V4, specialized for color perception, when processing speech or visual graphemes. Similarly, synesthetic tactile experiences engage the primary and secondary somatosensory areas more than controls. These studies suggest that synesthesia uses some of the same neural substrates that support veridical perception. Studies using fMRI have also directly contrasted synesthesia with color imagery and learned color associations and found that synesthesia is different from both. A recent event-related potential study of speech–color reports differences in amplitude in the EEG potential between synesthetes and controls at 122 milliseconds.

A number of other regions outside of the traditionally defined sensory cortex appear to be involved in synesthesia, particularly in the frontal and parietal regions. These may, in normal perception, be multisensory convergence zones and may be particularly involved in spatial aspects of perception. As such, their involvement in synesthesia is perhaps not surprising. Consistent with this, a recent fMRI study found activity in the posterior parietal cortex associated with number–space synesthesia.

Genes and Development

Synesthesia runs in families. Around 40% of synesthetes know of another first-degree relative with synesthesia, and there is little evidence for cultural transmission. Family members tend not to agree on the color of graphemes any more than unrelated synesthetes. Moreover, different types of synesthesia coexist within families. For example, synesthetes experiencing taste tend to have relatives with color synesthesia. This suggests that the gene has a restricted role in the determining the way that synesthesia manifests itself.

An examination of the variety of common inducer-concurrent pairings suggests a possible mechanism. Vilayanur Ramachandran and Ed Hubbard note that one reason why graphemes and colors may tend to be paired is because of their anatomical proximity within the fusiform gyrus within the visual ventral stream. This adjacency principle extends to other types of synesthesia, such as number–space and lexical–gustatory synesthesia. The gene may promote localized connectivity within the brain and, depending on where it is expressed, it may give rise to different forms of synesthesia.

One debate that has yet to be resolved is whether or not synesthesia reflects extra feedforward connections between two regions of the brain or removal of inhibition on normally dormant pathways. A study using diffusion tensor imaging, which demonstrates more localized organization of white matter in synesthetes, is perhaps more consistent with the former view. However, other phenomena, such as acquired forms of synesthesia, are compatible with a disinhibited feedback account. Acquired forms of synesthesia can occur after ingestion of certain drugs (e.g., LSD) and also as a result of sensory loss such as blindness or even blindfolding. Given that these symptoms can occur within hours (after drug use) or days (blindness), it is more consistent with a release from inhibition. It is, of course, conceivable that different causal mechanisms in developmental and acquired cases could produce similar perceptual outcomes.

Why Does Synesthesia Exist?

The fact that synesthesia is relatively common and has a genetic basis (in nonacquired forms) leads to the question of whether it could be evolutionarily adaptive. From first principles it need not be so. There are many traits and conditions with genetic components that are disadvantageous or just benign. Although there is little evidence that synesthetes are at any net disadvantage, it could possibly be construed a benign variant of normal perception. This is effectively a null hypothesis that one would fall back on in the absence of any evidence for an advantage. One postulated advantage of synesthesia is that it leads to creativity. The claim is that more widespread differences in brain connectivity could facilitate novel and adaptive ideas. Under this account, the perceptual characteristics of synesthesia are not directly relevant to synesthesia's true adaptive function. However, the evidence is mixed (synesthetes are more likely to be artists, but artists don't necessarily score higher on formal measures of creativity). Another candidate is that synesthesia may enhance memory because stimuli (e.g., words) may additionally be stored perceptually (e.g., as colors). There is strong evidence that synesthesia does lead to improved memory. Under this account, the unusual perceptual experiences would have a direct impact on this adaptive skill, perhaps accounting for its common existence.

Jamie Ward

See also Color Perception; Cross-Modal Transfer; Individual Differences in Perception; Multimodal Interactions: Neural Basis; Multimodal Interactions: Tactile–Auditory; Multimodal Interactions: Visual–Auditory

Further Readings

Banissy, M., & Ward, J. (2007). Mirror touch synaesthesia is linked with empathy. *Nature Neuroscience, 10,* 815–816.

Baron-Cohen, S., Harrison, J., Goldstein, L. H., & Wyke, M. (1993). Coloured speech perception: Is synaesthesia what happens when modularity breaks down? *Perception, 22,* 419–426.

Dixon, M. J., Smilek, D., & Merikle, P. M. (2004). Not all synaesthetes are created equal: Projector vs. associator synaesthetes. *Cognitive, Affective and Behavioral Neuroscience, 4,* 335–343.

Hubbard, E. M., & Ramachandran, V. S. (2005). Neurocognitive mechanisms of synesthesia. *Neuron, 48,* 509–520.

Nikolic, D., Lichti, P., & Singer, W. (2007). Color-opponency in synaesthetic experiences. *Psychological Science, 18,* 481–486.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia—A window into perception, thought and language. *Journal of Consciousness Studies, 8,* 3–34.

Simner, J., Mulvenna, C., Sagiv, N, Tsakanikos, E., Witherby, S. A., Fraser, C., et al. (2006). Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception, 35,* 1024–1033.

Ward, J. (2008). *The frog who croaked blue: Synesthesia and the mixing of the senses.* Oxford, UK: Routledge.

T

Tactile Acuity

When an object is too small or too far off in the distance, we cannot identify it visually. We can detect its presence but cannot make out its features. This inability to make out small features is caused by the limited spatial acuity (or spatial resolution) of the visual system. Our visual acuity is tested, using a so-called Snellen chart, by determining what the smallest identifiable letter is. Eye exams are, in effect, tests of how small visual features can be before we can no longer discern them.

The sense of touch is also subject to limited resolution. Indeed, when we run our fingers across surfaces with large features, such as the number pad on a phone, we can discern individual features (e.g., the buttons on the pad). On the other hand, if we scan a surface with small features, like that of an upholstered sofa, we discern texture but not individuated spatial features. This entry covers the tests of tactile spatial acuity, neural basis of tactile spatial acuity, the factors affecting spatial acuity, and why we should measure spatial acuity.

Tests of Tactile Spatial Acuity

Two-Point Threshold

The traditional test of tactile spatial acuity is the two-point threshold (2PT). The test consists of measuring how far apart two points contacting the skin need to be before they are perceived as two points rather than one. The 2PT is administered using a compasslike instrument in which the two rods of the compass have punctate tips. The test consists of pressing the two tips, separated by a predetermined amount, against the skin simultaneously and having subjects judge whether they feel one or two contact points. The test is repeated with a variety of distances between the two tips. The closer the two points are when the subject can still perceive them as two, the higher the acuity. The idea is that, to the extent that you can make out fine spatial features, you will be able to discern two tips as distinct when they are close together. If your tactile spatial resolution is coarse, you will perceive them as a single point.

Recently, however, the 2PT has fallen out of favor because it relies on subjective judgments: the perception of the two contact points falls along a continuum, from a single circular contact point (when the two contact points are adjacent or very close), to an increasingly elongated point (at a range of intermediate separations), to two points (at larger separations). The judgment of whether the two contact points feel distinct is therefore subjective.

One way to get around the subjectivity of the 2PT is to use a two-alternative forced choice task. On each trial, the subject is presented with two stimuli in sequence: One consists of a single tip and the other of two tips. Sometimes the single tip is presented first and the two tips second, sometimes the reverse is true. The subject's task is to judge which of the two stimuli (first or second) comprised two tips. In order to perform this task, the subject

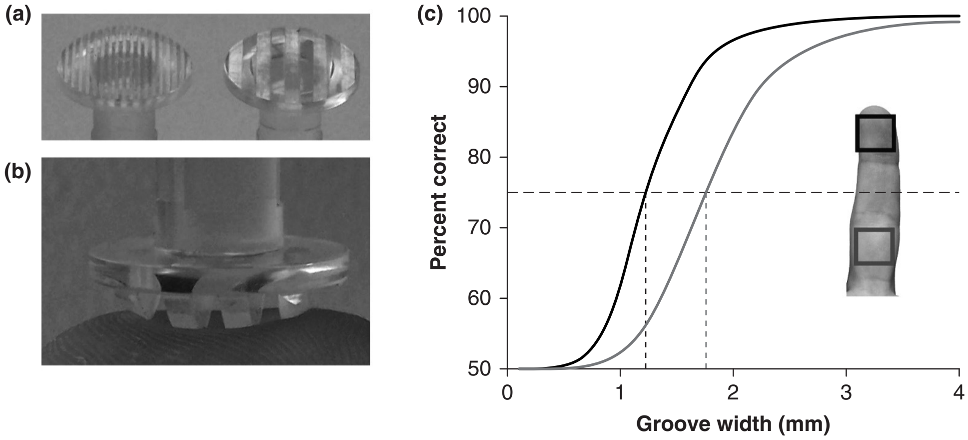


Figure 1 Measuring the Grating Orientation Threshold

Notes: (a) Two JVP domes, one with a groove width of 0.5 mm (left), the other with a groove width of 2 mm (right). The first is below GOT, the other above. (b) Grating pressed against the skin of the finger. The subject's task is to identify the orientation of the grating relative to the orientation of the finger. (c) GO task performance as a function of the groove width for two locations on the finger, the fingertip (black box) and the base of the finger (gray box). Spatial acuity is higher for the fingertip than the finger base, as evidenced by the fact that the psychometric function of the former is shifted left relative to that of the latter. The gray horizontal dashed line shows the criterion of performance for the determination of the GOT. The GOT is 1.25 mm for the fingertip and 1.75 mm for the finger base.

has to be able to distinguish one from the other. However, this task is problematic in that nonspatial cues may help subjects perform it. For instance, two nearby tips may feel like a single contact point with a dull tip whereas the single tip feels sharper, a judgment based on an intensive rather than spatial difference.

Grating Orientation Discrimination

The 2PT has been replaced by a more rigorous and reliable test, the grating orientation (GO) discrimination test. In this test of tactile spatial acuity, a tactile grating—consisting of alternating ridges and grooves, typically of equal width—is pressed against the skin (Figure 1). The grooves and ridges are either parallel to the long axis of the finger (vertical) or perpendicular to it (horizontal). The subject's task is to determine the orientation of the grating (vertical or horizontal). In order to perform this task above chance, the subject must perceive the spatial configuration of the grating in a manner analogous to perceiving the spatial configuration of a letter in a test of visual acuity. For fine gratings, with grooves narrower than about 1 millimeter (mm), subjects are nearly at chance (50% correct) on the GO discrimination (Figure 1). As the width of the grooves increases, however, performance improves. The grating orientation threshold (GOT) is usually the groove width at which subjects perform the task with 75% accuracy (Figure 1). The GOT can then be compared across individuals or across body regions. Unlike the 2PT, the GO task does not have a subjective component.

Spatial Localization

Another way in which tactile spatial acuity can be measured is by determining how well a subject can identify where he or she has been touched. The test involves applying a rod with a small punctate tip (similar to the one used in the single contact condition of the two-alternative force choice 2PT measurement) and having the subject identify the location at which he or she was touched as accurately as possible. The distance is then measured between the location at which the subject was touched and the location at which he or she reported being touched. The higher the spatial acuity, the smaller this distance is. In general, spatial localization is better than 2PT or GOT.

Neural Basis of Tactile Spatial Acuity

All tactile sensations begin as a pattern of neural activity in nerve fibers whose signals originate from receptors in the skin. There are a variety of receptors in the glabrous (hairless) skin, but only two types of receptors convey information about the spatial configuration of stimuli impinging upon it—Merkel's disks and Meissner's corpuscles—both of which respond to deformations of the skin.

Signals from each of these two mechanoreceptors are conveyed to the brain by distinct nerve fibers. The spatial layout of a stimulus delivered to the skin is reflected in the spatial pattern of activity it evokes in Merkel's and Meissner's receptors. For instance, when a tactile grating is pressed onto the skin, mechanoreceptors located under the ridges of the grating will be more excited than mechanoreceptors located under the grooves because the skin deformation is greater under the ridges than it is under the grooves. Our ability to perceive the grating relies on this pattern of mechanoreceptor activity. When the grooves of the grating are large, the spatial pattern of activity in mechanoreceptors will consist of bands of very active mechanoreceptors interleaved with bands of inactive mechanoreceptors. Because the distance between adjacent mechanoreceptors in the skin of the fingertip is approximately 1 mm, a fine grating (i.e., one with grooves less than 1-mm wide) may not yield the required pattern of neural activity. Indeed, in order to perceive the orientation of a tactile grating, one row or column of active receptors must be interleaved with another inactive row or column of receptors. For that to be the case, the groove must be at least 1 mm wide. The same reasoning holds for 2PT: The mechanoreceptors excited by the two tips must be separated by one silent mechanoreceptor.

In touch as in vision, then, spatial acuity is, for the most part, limited by the density of receptors in the retina or in the skin. In both modalities, the GOT is approximately equal to the separation between two adjacent receptors. Therefore, although our visual acuity is better than its tactile counterpart in that we can identify a smaller letter by vision than by touch, the two are equivalent in that, if a tactile letter and a visual letter excite the same number of receptors (mechanoreceptors and photoreceptors, respectively), our ability to identify that letter will be equivalent in the two modalities.

Importantly, regions with higher acuity, such as the fingers or lips, are ascribed larger representations in the brain. Indeed, there are approximately as many neurons in the brain that receive information from the lips as there are neurons that receive signals from the shoulder, trunk, hips, and thighs. Without this increased representation in the brain, referred to as *cortical magnification,* changes in receptor density in the skin would not result in the substantial differences in spatial acuity observed. Enough brain volume has to be dedicated to signals emanating from highly acute body regions to be able to process the fine spatial information they convey.

Factors Affecting Spatial Acuity

Tactile spatial acuity varies considerably across the body surface, as visual spatial acuity has been shown to vary across the visual field. Indeed, in vision, spatial acuity is best in the center and deteriorates at the periphery of the visual field. In touch, the most sensitive regions are the lips and fingertips, where the GOT can be less than 1 mm. The back, legs, shoulders, and arms yield much higher thresholds, measured in the tens of millimeters. Differences in spatial acuity reflect differences in the sensory functions carried out with the different regions of the body surface. For instance, the tactual exploration and manipulation of objects is typically performed with the hand. These functions require that we are able to perceive the object we are palpating or manipulating. In contrast, we do not explore or manipulate objects with our backs, so being able to sense the fine structure of objects with this body region is not important. The lips serve at least two important functions for which acute sensation plays a crucial role: talking and kissing.

Also, spatial acuity declines gradually with age, from 1 mm (on the fingertip) for young adults to over 2.5 mm in some subjects over 60 years old. This decrement in spatial acuity with age is not due to changes in the properties of the skin itself but rather to a decrease in the number of receptors embedded in the skin and to a deterioration of those receptors that remain.

A controversial question in the study of tactile spatial acuity is whether blindness leads to an increase in tactile acuity. Indeed, visually impaired subjects rely more heavily on their sense of touch than do sighted individuals, and it has been shown that brain areas that process visual information in sighted people respond to tactile stimulation in blind people. An intriguing possibility is that this increased brain area responding to touch may lead to improvements in tactile acuity. Recent evidence suggests that blind people may have slightly but significantly lower GOTs than their sighted counterparts. Because it is unlikely that the skin of blind people comprises more mechanoreceptors than does that of their sighted counterparts, the superior spatial acuity in blindness is likely due to modifications in the brain, possibly to the increased brain volume that responds to tactile stimulation. This difference between blind and sighted subjects is important because it shows that the heightened reliance on touch, with the concomitant increase in tactile experience, can lead to changes in how the brain processes tactile information. The superior acuity in blindness is also an important result because it suggests that spatial acuity is not entirely determined by limitations at the sensory periphery.

Why Measure Spatial Acuity?

Measures of spatial acuity constitute an important component in the assessment of the perceptual capacities of individuals with nerve or brain damage or other potentially impairing conditions. In the case of vision, the conditions need not be severe. Indeed, measures of visual spatial acuity are most often used to determine what type of optical correction individuals need to achieve normal vision (e.g., for fitting them with glasses or contact lenses). Importantly, visual spatial acuity is a reliable gauge of how good vision is in general: Individuals with good vision have good spatial acuity (and, typically, vice versa).

Measures of tactile spatial acuity are less common because they are typically only called upon when the nerve or brain is liable to have been damaged. For instance, a patient may report abnormal sensation in a body region. If his or her tactile spatial acuity is impaired (i.e., if the GOT is substantially higher than the average for that body region), the patient is likely to have incurred some damage along the somatosensory pathway. The physician can then seek to identify the locus of the damage, that is, whether it be damage of the skin (and thus mechanoreceptors embedded in the skin), of the nerve that carries signals from mechanoreceptors to the brain, or of the brain itself. Whatever the case may be, tests of spatial acuity are critical diagnostic tools in identifying perceptual impairments in the first place.

Sliman Bensmaia

See also Ageing and Touch; Body Perception; Braille; Cortical Reorganization Following Damage; Cutaneous Perception; Cutaneous Perception: Physiology; Haptics; Neural Representation/Coding; Texture Perception: Tactile; Virtual Reality: Touch/Haptics; Visual Acuity; Visual Disorders: Blindness; Visual Processing: Retinal; Visual Receptors and Transduction

Further Readings

Craig, J. C., & Johnson, K. O. (2000). The two-point threshold: Not a measure of tactile spatial resolution. *Current Directions in Psychological Science, 9,* 29–32.

Goldreich, D., & Kanics, I. M. (2003). Tactile acuity is enhanced in blindness. *Journal of Neuroscience, 23,* 3439–3445.

Johnson, K. O., & Phillips, J. R. (1981). Tactile spatial resolution: I. Two-point discrimination, gap detection, grating resolution, and letter recognition. *Journal of Neurophysiology, 46,* 1177–1191.

Stevens, J. C., & Patterson, M. Q. (1995). Dimensions of spatial acuity in the touch sense: Changes over the life span. *Somatosensory & Motor Research, 12,* 29–47.

Van Boven, R. W., & Johnson, K. O. (1994). A psychophysical study of the mechanisms of sensory recovery following nerve injury in humans. *Brain, 117,* 149–167.

Wheat, H. E., Goodwin, A. W., & Browning, A. S. (1995). Tactile resolution: Peripheral neural mechanisms underlying the human capacity to determine positions of objects contacting the fingerpad. *Journal of Neuroscience, 15,* 5582–5595.

TACTILE MAP READING

Extracting meaningful information from a tactile map, that is a map elevated in the third dimension, designed to be read by the sense of touch, is far more problematic than reading a conventional map with the use of vision. Tactile maps are widely used in educational settings and in orientation and mobility training for visually impaired individuals. Maps and graphics are the most fundamental and primary mechanism for communicating spatial arrangements to blind people—that is, any representation of spatial features, their arrangement, and intrarelationships. Tactile graphics are used as diagrams in school textbooks and portable maps when traveling. Just as Braille is often used as a substitute for the written word, tactile graphics are the equivalent for maps and diagrams. These are essential tools for providing independence and education to people without vision. This entry describes the limitations of tactile map reading and the strategies to overcome them.

Limitations

The first significant problem is that a tactile map has to be designed to be read by the sense of touch, consequently, the map can only be read by the resolution of tactile apprehension of the fingertips. Although it is possible to sense very small areas with the fingertips (for example, the point of a pen), for a map to be useful, an often quoted general design criteria is that the working area that can be sensed at any one time is the area under a fingertip. Therefore, the resolution of a tactile map is approximately the area under a fingertip, approximately 1 centimeter (cm) by 1 cm. This presents the potential map reader with a significant problem in that the amount of information that he or she can gain from a map is limited to chunks of information of a fingertip in size.

Perhaps a more significant problem than the limitations of the individual area that can be sensed any one time is that these areas are sensed in isolation, as the fingertip sequentially moves from one area to another. With vision it is possible to gather a Gestalt-like overview of a visually read map in its entirety and the relationships between all of the features. However, when reading a tactile map, the map reader has to move his or her fingers over the surface of the map sensing simplified and enlarged features and then has to piece these individual components of information together to generate a holistic overview of the map. If a conventional visual map was rendered tactile, it would be unreadable by the fingertips due the density, clutter, and complexity of information. An analogy to tactile map reading is to imagine that you have to read, integrate, and comprehend all of the features on a map; however, the map is obscured by a layer of paper. You are then only able to view the map using vision through a hole in the paper, which is the size of a fingertip, and by moving this small window around the map. This provides a general idea as to the problems presented by tactile map reading.

Cartographic representational issues are greatly exaggerated by the nature of this fingertip reading. Any text labeling has to be undertaken in Braille. Individual Braille cells are 4 by 6 millimeters (mm). Therefore, large areas of any tactile map may be covered in Braille labels. Production methods for tactile maps include hand- or homemade techniques, thermoform, vacuum pressing, microcapsule paper, and, more recently, inkjet printing and specialized raised line printers and embossers.

Perhaps the biggest problem with tactile map reading is that a tactile map, while very good at providing an overview of the spatial arrangement of features through its tangible surface that can be explored at the leisure of the user, cannot be queried for more context or deeper levels of information. This information would be readily available to a visual map reader viewing the full compliment of information on a map through other mechanisms, such as color, legends, text, and so forth. The tactile map reader needs to be provided with information about the content of the map, what is displayed on the map, and how it is symbolized. This is far more necessary than with visual information. As an exaggerated example, a circle rendered on a tactile graphic without any contextual information, such as labeling, is initially difficult for the tactile map reader to independently gather the necessary information to know what the circle is: a globe, a geometric circle, the top of a coffee cup, or is the circle representing a symbol for a city on the map? For tactile map reading to be effective there has to be an increased level of map awareness and context when the map is explored by the user. When an individual visually reads a map or graphic, this information is readily apparent. The map reading task for the blind individual is a sequence of becoming oriented to the information on the map, or at least acquiring the knowledge of what the map is representing, legend comprehension, map and legend integration, and then information extraction. The ability to simultaneously read a map and a legend at the same time it is also problematic, as the need to refer to a legend during map reading disrupts the referencing task and the spatial exploration of the map.

Strategies to Overcome Limitations

In spite of these major limitations, users without vision are able to gain meaningful information from tactile maps and to integrate this into maplike representations. Their ability to do this is often related to their tactile map reading strategies. These strategies include a two-handed strategy, where one hand serves as a spatial anchor on the map and the other roams across the map surface. One-handed scanning techniques include ordered scanning, from left to right, up and down, and circular motions moving around the map and then gradually focusing in on the center; other users adopt a quick overview by roughly scanning the map to try and gather the most salient features, where they are located, and then return to explore these in more detail.

To overcome many of the problems with tactile maps, researchers have investigated ways of rendering them in the digital domain, using touch pad technology for annotating conventional tactile maps with speech labels to bypass Braille labeling issues. Other approaches have presented audio information or levels of information using computers and interactive methods through Web pages. Using specialized human–computer interface equipment, maps may be disseminated across the Internet and touched and heard by a user in another location by using haptic devices. A haptic device is a computer mouse that moves in three-dimensional space, providing resistance and varying feedback to the user so that they are able to feel a maplike surface.

Irrespective of the production method or the technology used, the central issue for the tactile map reader is the complex scenario of building an integrated understanding of a map scene via searching with the hand and fingertips.

R. Daniel Jacobson

See also Braille; Cutaneous Perception; Haptics; Kinesthesia; Proprioception; Spatial Memory; Tactile Acuity

Further Readings

Blades, M., Ungar, S., & Spencer, C. (1999). Map use by adults with visual impairments. *The Professional Geographer, 51*(4), 539–553.

Golledge, R. G., Rice, M., & Jacobson, R. D. (2005). A commentary on the use of touch for accessing on-screen spatial representations: The process of experiencing haptic maps and graphics. *The Professional Geographer, 57*(3), 339–349.

Jacobson, R. D. (1998, August). Navigating maps with little or no sight: A novel audio-tactile approach. *Proceedings of Content Visualization and Intermedia Representations*. University of Montreal, Canada.

Jacobson, R. D., & Kitchin, R. M. (1997). GIS and people with visual impairments or blindness: Exploring the potential for education, orientation, and navigation. *Transactions in Geographic Information Systems, 2*(4), 315–332.

Ungar, S., Blades, M., & Spencer, C. (1997). Strategies for knowledge acquisition from cartographic maps by blind and visually impaired adults. *Cartographic Journal, 34*(2), 93–110.

TASTE

Taste refers to the sensory system in which water-soluble chemosensory stimuli are converted into neural signals and perceptual experience. Although taste perception is essential to enjoying a well-prepared steak or lasagna, it also allows us to survive by assisting in the avoidance of toxins and in the regulation of energy balance (homeostasis). In fact, the primary function of taste is to aid in the ingestion of nutrients and in the avoidance of noxious compounds. For example, the energy-rich fructose of an apple tastes sweet and pleasurable, yet the poisonous alkaloid metabolites in plants such as hemlock taste bitter and aversive.

Taste perception begins in specialized structures in the tongue called taste buds that send signals to the brain through peripheral nerves. The peripheral and central taste pathways enable the perception of thousands of different chemosensory stimuli, each evoking a basic taste quality: sweet, sour, salty, bitter, and perhaps *umami* (savory). Taste quality–specific information is transduced by dedicated receptor mechanisms at the periphery before converging onto cells in the central nervous system. In the brain, taste information is processed with respect to previous experience, emotional salience, and nutritive value. This multimodal processing informs ingestive behavior and enhances survival. The topics covered in this entry are taste and nutrition, taste stimuli and quality, taste transduction, peripheral and central taste anatomy, and the neural coding of taste.

Taste and Nutrition

In addition to discerning food from poisons, taste perception allows us to gather nutritional information about the substances we ingest and to weigh this information against the body's homeostatic needs. Salty foods that provide ions critical for cell function are distinguished from sweet foods that provide energy-rich carbohydrates. Even nutrients such as fat and water have been shown to evoke tastes. The taste of food is acutely modified by our internal or homeostatic state. The concentration of critical nutrients in the body is monitored continually, and a change in the homeostatic state can cause alterations in the palatability, or pleasantness, of particular tastes. For example, in conditions of extreme salt deprivation, the body compensates by allowing extremely salty and normally aversive foods to taste good. As a consequence, the organism will consume a larger than normal quantity of salt, and homeostasis will be restored. The adaptability of the gustatory system informs ingestive behavior and is critical for maintaining a homeostatic balance of nutrients in the body.

Taste Stimuli and Quality

Taste stimuli are called *tastants* and can be characterized by the perceptual "quality" they evoke. In human taste perception, there are four basic taste qualities: sweet, salty, sour, and bitter. Each taste quality is converted into a neural signal by a dedicated receptor mechanism at the periphery. The argument for the existence of a fifth basic taste quality, called umami (a Japanese word for savory), has been bolstered by recent empirical data suggesting that umami substances activate a discrete receptor mechanism on the tongue. Chemosensory stimuli that evoke a particular taste quality are often but not always of similar chemical composition. In some cases, the taste quality evoked by a particular stimulus depends on its concentration. For example, some salts will taste sweet at low concentrations.

The sensation of sweetness is most commonly evoked by sugars, such as sucrose, maltose, dextrose, and fructose. Sugars constitute carbohydrates or saccharides, one of the four major macromolecules. Saccharides are metabolized quickly and are thus an important source of energy. The pleasantness of sweet stimuli therefore encourages the consumption of energy-rich carbohydrates. Other substances that evoke sweetness include proteins and their constitutive amino acids, low concentrations of some salts, and polyalcohols.

Saltiness is the taste quality most commonly associated with the salt sodium chloride. Salts provide ions that are critical for many cellular functions, including neural signaling and osmotic balance. Due to its essential role in homeostatic regulation, salt is appetitive to most mammals at low to moderate concentrations. Other compounds that can evoke saltiness include lithium chloride and sodium sulfate.

Sourness is most commonly evoked by acids such as hydrochloric acid and citric acid. In nature, acidic compounds are associated with both unripened fruits and rotten foods. Sour stimuli therefore signal avoidance in most mammals. Other examples of sour-inducing acids include tartaric acid (found in fruits) and lactic acid (produced by bacteria in the process of fermentation).

Bitterness is a taste quality most commonly evoked by the plant alkaloid quinine. Other substances that evoke bitterness include urea, nicotine, caffeine, and some salts. Bitterness is associated with toxicity, as many plant toxins are alkaline in nature and thus evoke a strong and aversive bitter sensation. The perception of bitterness aids in the avoidance of these noxious substances.

Umami is the controversial fifth taste quality that is evoked by the chemicals monosodium glutamate (MSG), 5'-inosate (IMP), and disodium 5'-guanosalte (GMP). These chemical compounds are common to a variety of meats and cheeses and can also be found in fish, tomatoes, and some varieties of mushroom. Umami is generally a hedonically positive taste quality that signals food acceptance. The controversy as to whether it is a taste quality unto itself stems from a dearth of evidence of cells in the central nervous system that respond specifically to umami.

Taste Transduction

Transduction is the process through which environmental stimuli are converted into neural signals. Environmental stimuli can take many forms, such as photic or light stimuli for vision, mechanical or thermal stimuli for touch, and chemical stimuli for taste and smell. Each sensory system has evolved dedicated receptor mechanisms that are sensitive to a particular type of stimulus. Transduction produces neural signals in the form of nerve impulses called *action potentials* or *spikes*. Gustatory transduction occurs at specialized epithelial cells called taste receptor cells (TRCs). TRCs regenerate approximately every 10 days and can be found in groups of 50 to 100 in clusters called taste buds.

Taste buds are distributed throughout the tongue and oropharyngeal cavity, including the soft palate, epiglottis, larynx, and pharynx. In addition to TRCs, taste buds contain basal cells that are a form of stem cell and supporting cells, which aid in TRC function. Taste buds of the tongue are situated on protrusions called papillae. The chemoreceptive tissue of the tongue is divided topographically into three morphologically distinct forms of papillae. Papillae of the anterior two-thirds of the tongue are referred to as *fungiform,* whereas papillae of the posterior one-third of the tongue are *circumvallate,* and papillae lining the lateral edges of the tongue are called *foliate*.

Taste receptor cells transduce neural signals in a stimulus-dependent manner. There are two general mechanisms through which a chemosensory stimulus can activate a TRC. The first mechanism involves the direct passage of ions through a channel by a chemosensory stimulus. This is the general mechanism for the transduction of salty and sour stimuli. In the transduction of salts, for example, the positively charged sodium ions dissociate from the negatively charged chloride ions and flow into the cell membrane through dedicated sodium channels. In the second mechanism of taste transduction, a chemosensory stimulus will bind to a specific G protein coupled receptor (GPCR) on the surface of the cell. The GPCR is embedded in the cell membrane and, when activated by the presence of a taste stimulus, serves to initiate a process of intracellular communication called a *signaling cascade*. The signaling cascade entails a chain of communication within the cell through chemicals called *second messengers* that function in the opening of TRC ion channels. The more circuitous second mechanism is used in the transduction of sweet and bitter stimuli and stimuli evoking umami. The second-messenger system relies on two families of receptor proteins called the T1Rs and the T2Rs. Combinations of the various receptor subtypes provide quality specific mechanisms of taste transduction.

Each taste quality is transduced by its own receptor mechanism(s). The transduction of saltiness involves the passage of sodium ions directly into the TRC through two different types of sodium channels. The first type of sodium channel can be blocked by the diuretic amiloride and is located on the apical portion of the TRC. The second type of sodium channel is insensitive to amiloride and is located on the basolateral membrane of the TRC. The passage of sodium ions through each of these channels will result in the opening of several other ion channel types and will facilitate the signaling of saltiness. There are several proposed mechanisms of sour transduction that differ depending on the species of interest. The passage of protons directly through a hydrogen ion channel has been proposed, however, the channel has not yet been identified conclusively. Other proposed mechanisms involve activation of stimulus-gated calcium channels and amiloride-sensitive sodium channels.

Sweet transduction relies on a second-messenger signaling cascade that begins when a sweet stimulus binds to a particular receptor. The transduction of sweet stimuli relies on two GPCRs, called T1R2 and T1R3. Taste receptor cells that transduce sweet stimuli express a combination of the T1R2 and T1R3 receptors. The transduction of bitter stimuli also depends on second-messenger cascades initiated by GPCRs, specifically the T2R receptors. There are many subtypes of the bitter receptor, and therefore a diverse array of taste stimuli elicit bitterness. Umami transduction also relies on second-messenger cascades initiated by a combination of the T1R1 and T1R3 receptors. In rodents, the T1R3 receptor is responsive to L-amino acids; however, in humans, the receptor binds only umami-inducing compounds.

Peripheral and Central Taste Anatomy

There are three gustatory nerves that innervate the tongue in a regionally specific manner. The chorda tympani is a branch of the facial nerve (cranial nerve VII) that innervates the fungiform papillae of the anterior two-thirds of the tongue. This nerve is essential for taste discrimination and particularly in the detection of salts. The greater superficial petrosal is also branch of the facial nerve that innervates taste cells of the palate. The glossopharyngeal nerve (cranial nerve IX) innervates taste cells of the caudal one-third of the tongue. The glossopharyngeal has been shown to play an essential role in orofacial reflexive behavior associated with taste responsivity, particularly with respect to aversive stimuli. The superior laryngeal branch of the vagus nerve (cranial nerve X) innervates the larynx and epiglottis. The superior laryngeal is essential to some aspects of ingestive behavior, such as swallowing.

Taste information from cranial nerves VII, IX, and X arrives in the central nervous system through projections to cells in the rostral subdivision of the nucleus tractus solitarius (NTS) in the caudal medulla. The NTS is the first central relay for taste processing. The peripheral taste nerves make excitatory connections on the cells of the NTS; however, inhibitory interneurons are quite numerous in this structure. In most mammals, cells of the NTS project to the parabrachial nucleus of the pons and then to the ventromedial thalamus. In many primates, including humans, the NTS bypasses the parabrachial pons and projects directly to the thalamus. From the thalamus, projections extend to the gustatory cortex located in the insula.

In addition to the "main" gustatory pathway, there are many structures that either receive gustatory information or provide input to the main gustatory nuclei. This reflects the important role of gustation in ingestive behavior, homeostatic regulation, and reward. Structures that are important to gustatory processing include the reticular formation, lateral hypothalamus, central nucleus of the amygdala, bed nucleus of the stria terminalis, and the substantia innominata. Each structure is involved in a specific taste-related function. Reciprocal connections providing a descending modulatory influence include projections from the gustatory cortex, lateral hypothalamus, and central nucleus of the amygdala. Each contributes to the initial processing of taste in the brain stem. Figure 1 provides a graphic overview of the anatomy of the taste system.

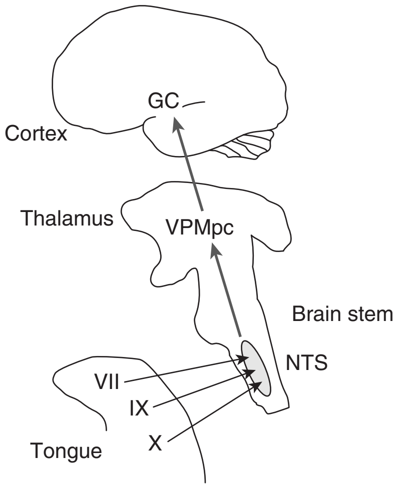


Figure 1 A Diagram of the Anatomical Pathways Associated With Taste in the Human Brain

Notes: Abbreviations are as follows: VII = facial nerve; IX = glossopharyngeal nerve; X = vagus nerve; NTS = nucleus of the solitary tract; VPMpc = ventroposteromedial nucleus of the thalamus, parvocellular region; GC = gustatory cortex.

Neural Coding of Taste

Neural coding refers to the representation of environmental stimuli with neural activity. When a taste stimulus is bathed over the tongue, taste-responsive neural elements (i.e., either peripheral nerve fibers or neurons in the brain) will most often increase their rate of firing. Though different cells have different sensitivities to the various taste qualities, most cells respond to more than one taste quality. In general, changes in stimulus concentration (intensity) result in changes in firing rate: higher concentrations evoke higher firing rates. As in other sensory systems, each log unit increase in stimulus concentration will result in a linear increase in firing rate up to a saturation point, after which no further increase in firing rate is apparent. The slopes of these intensity–response relationships vary across cells. In studies of responses to taste mixtures, both mixture suppression (the response to the mixtures are less than the sum of the responses to the components) and mixture enhancement (responses to mixtures exceed the sum of the responses to the components) have been noted. Most often, however, the response to a taste mixture is closely predicted by the response to the component of the mixture that evokes the largest response when presented alone.

There are several ways in which a stimulus can be represented in the brain. For example, different taste qualities might be signaled by different neurons. In that case, each taste quality would be associated with a separate subset of neurons. This is called *labeled line* coding. There is good evidence that peripheral nerve fibers that innervate the tongue use this method of coding. That is, there appear to be dedicated channels of information concerned with each taste quality that, when eliminated through genetic deletion of specific types of taste receptor proteins, largely deprive the animal of the ability to perceive that quality. Another strategy that can be used by neurons to represent the taste world is through the distribution of activity across the population of neurons. So, in that case, a distinctive "across neuron pattern" of activity would be evoked by each taste quality. In the brain stem, where there is abundant convergence of different types of peripheral nerve fibers onto single cells, some evidence for both labeled line and across neuron pattern coding has been reported.

The origin of the labeled line theory as it is applied to gustation is based on the observation that, although most taste-responsive cells at all levels of the nervous system respond to more than one of the basic taste qualities, knowledge of the quality to which a cell responds most vigorously, called its "best" stimulus, allows one to predict the relative sensitivity of that cell to the other taste qualities. For example, a salt best cell will most often respond next-best to sour, and then to bitter; a sweet best cell may also show some sensitivity to salt, and so on. More recently, the classification scheme that many investigators use has shifted from a strictly best stimulus categorization, with each best stimulus cell type associated with a single taste quality, to a scheme of "specialists" versus "generalists." In the latter grouping, specialists are most narrowly tuned to respond to a single taste quality, their best stimulus, whereas generalists respond more broadly across taste qualities.

Functional evidence for labeled lines in the taste system derives from experiments that correlate behaviors related to a specific taste quality with changes in the subset of neurons that respond best to that taste quality. For example, sodium deprivation, which results in an increased preference and appetite for very concentrated salt, will change the responses to NaCl only in NaCl best cells. Conversely, application of amiloride, a sodium channel blocker, to the tongue blocks responses to salt mostly in salt best cells. Another example comes from experiments where animals are trained to avoid sweet saccharin by pairing it with experimentally induced illness. In that case, responses to sweet substances are altered only in sweet best cells. Changes in the relative response rates across tastants, called a cell's *response profile,* that are induced by dietary or experiential changes illustrate the idea that responses across taste qualities in a given cell are adaptive and malleable. In fact, it has been shown that brief pulses of a taste stimulus just prior to presentation of a second different taste stimulus can change the response profile of cell in the brain stem, suggesting that even the immediate taste history or context can affect the response properties of taste cells.

An alternative to the labeled line theory, but not incompatible with it, is the across fiber or across neuron pattern theory. This theory is based on the observation that similar tasting stimuli will evoke similar patterns of activity across a population of cells, whereas dissimilar tasting stimuli will evoke dissimilar patterns of activity. When experimental manipulations, such as conditioned taste aversion (where a novel taste stimulus is rendered aversive by pairing it with illness) or sodium deprivation, alter responses to specific tastes in subsets of cells as previously described, they also change the across neuron patterns associated with those tastes. However, there are some conditions, (e.g., conditioned taste preference, where a novel taste stimulus becomes preferred by pairing it with a reward) that only affect the across neuron pattern and do not target any specific cell type. Importantly, similarity and differences among across neuron patterns in the brain stem evoked by a variety of taste stimuli bear a close relationship to behavioral reactions to these stimuli.

In spite of several decades of debate among theorists, the relative merits of the across neuron pattern theory compared with the labeled line theory are still the subject of controversy. Both of the labeled line and across neuron pattern theories of taste coding rely on the ability of the nervous system to integrate neural activity over some interval of time. That is, the basic measure of a taste response is the firing rate over a period of time that includes part or all of the taste stimulus presentation, essentially ignoring the temporal characteristics of the response. However, changes in firing pattern of the time course of the response may represent another coding strategy available to the nervous system.

This idea, called *temporal coding,* actually encompasses a few different coding mechanisms, all of which have in common the fact that the timing of neural activity in one or more taste-responsive neurons represents the taste stimulus. For example, systematic changes in the firing rate over the course of the response, also known as the *time course* or *rate envelope* of the response, are one mode of temporal coding. Another is the timing of nerve impulses with respect to the onset of the taste stimulus. One or both of these mechanisms have been shown to convey information about the taste quality of the stimulus in the brain stem, thalamus, and gustatory cortex. A third temporal coding mechanism is that of synchronous firing among groups of neurons. In the gustatory cortex, for example, it has been shown that a given neuron may synchronize its firing with different groups of neurons, depending on which stimulus is presented. The existence and significance of these "dynamic assemblies" of neurons is just beginning to be explored by investigators.

Just how taste stimuli are encoded by the nervous system remains a topic of active investigation in the field. There is evidence for the use of labeled lines across neuron patterns and temporal coding from the periphery to the cortex. That said, it is certainly possible, if not probable, that all of these coding strategies are used to encode the various nuances of a taste stimulus, including its quality, its concentration, and its pleasantness or unpleasantness. Moreover, it is also possible that different coding mechanisms signal different aspects of a taste stimulus. Resolution of these issues awaits further investigation.

Patricia M. Di Lorenzo and Andrew M. Rosen

See also Taste: Disorders; Taste: Genetics of; Taste: Supertasters; Taste Adaptation; Taste and Food Preferences; Taste Receptors and Transduction; Taste Stimuli: Chemical and Food; Taste System Structure; Taste Thresholds and Intensity

Further Readings

Breslin, P. A., & Huang, L. (2006). Human taste: Peripheral anatomy, taste transduction, and coding. *Advances in Oto-Rhino-Laryngology, 63,* 152–190.

Chandrashekar, J., Hoon, M. A., Ryba, N. J. P., & Zuker, C. S. (2006). The receptors and cells for mammalian taste. *Nature, 444,* 288–294.

Di Lorenzo, P. M. (1999). The neural code for taste in the brain stem: Response profiles. *Physiology and Behavior, 69,* 87–96.

Hallock, R. M., & Di Lorenzo, P. M. (2006). Temporal coding in the gustatory system. *Neuroscience and Biobehavioral Reviews, 30*(8), 1145–1160.

Katz, D. B., Nicolelis, M. A. L., & Simon, S. A. (2002). Gustatory processing is dynamic and distributed. *Current Opinion in Neurobiology, 12,* 448–454.

Scott, T. R., & Giza, B. K. (2000). Issues of gustatory neural coding: Where they stand today. *Physiology and Behavior, 69,* 65–76.

Smith, D. V., & St John, S. J. (1999). Neural coding of gustatory information. *Current Opinion in Neurobiology, 9,* 427–435.

Smith, D. V., St John, S. J., & Boughter, J. D., Jr. (2000). Neuronal cell types and taste quality coding. *Physiology and Behavior, 69,* 77–85.

Spector, A. C., & Travers, S. P. (2005). The representation of taste quality in the mammalian nervous system. *Behavioral and Cognitive Neuroscience Reviews, 4*(3), 143–191.

TASTE: DISORDERS

Taste disorders are broadly classified as losses (i.e., ageusia and hypogeusia) or as distortions of gustatory function (i.e., dysgeusia). Ageusia is the complete absence of the ability to detect or recognize any taste quality, including sweet, sour, salty, bitter, as well as less familiar taste sensations, including umami (meatlike taste of glutamate salts), fatty, metallic, starchy/polysaccharide, chalky, and astringent. Hypogeusia is the reduced ability to detect or recognize taste stimuli and refers both to generalized loss of sensitivity across all taste qualities or to decrements that are limited to a specific taste quality, such as salty or bitter. A person with hypogeusia requires more molecules (or ions) for a taste sensation to be perceived and recognized compared to a person with normal taste perception.

Dysgeusia may or may not be correlated with loss of sensitivity and can occur in the presence or absence of a stimulus. To a person with dysgeusia, the presence of foods or beverages in the mouth that are normally pleasant may taste unpleasant or induce an inappropriate taste sensation, such as metallic. Furthermore, the time course of taste sensations may be altered, resulting in taste perseveration (prolonged persistence) or rapid extinction. Abnormal taste sensations may even occur when food and beverages are not present in the oral cavity. Loss or diminution of sensitivity for a single taste quality, such as salty, can distort the overall taste of a mixture resulting in dysgeusia. Most patients with taste disorders exhibit hypogeusia or dysgeusia; ageusia is rare. People with smell disorders often mistake their symptoms for a taste disorder. This confusion arises because odor sensations typically occur simultaneously with taste signals during mastication of food. Odorous molecules released from food in the oral cavity are transported to smell receptors up the back of the throat to the nasal cavity (called retronasal olfaction). Reductions in odor sensations via retronasal olfaction simply reduce the odorous component of the flavor of food but not the gustatory (taste) component. This entry describes the occurrences and causes of taste disorders, as well as burning mouth syndrome and recovery and treatment of taste disorders.

Occurrence

Taste disorders occur with greater frequency in older individuals and are exacerbated by certain medical conditions, pharmacologic and surgical interventions, radiation, and environmental exposure to toxic chemicals. Because the prevalence of taste disorders tends to increase with age, the number of people afflicted with taste disorders will escalate over the next decades as a result of the exponential growth in the older segment of the global population as projected by the United Nations. Taste disorders experienced by older people in the absence of diseases, medications, and other confounding variables is predominantly due to the fact that age-related losses are not uniform across compounds, but rather are dependent on the chemical structure of the tastant. For sodium salts, losses at the threshold level are greater for sodium sulfate and sodium tartrate than sodium chloride (table salt). Losses in the perceived intensity of the amino acids L-aspartic acid and L-glutamic acid are far greater than for other amino acids ,such as L-lysine and L-proline. For sweeteners, large molecules such as thaumatin, rebaudioside, and neohesperidin dihydrochalcone exhibit greater losses than sweeteners with lower molecular weights, such as sucrose and glucose. The relative differences in loss for individual compounds with age induces taste distortions (dysgeusia) because foods and beverages are comprised of mixtures of many different chemical compounds (i.e., a mixture will taste different to older people than younger ones due to age-related differences in the relative sensory salience of individual compounds).

Taste disorders have been reported by patients with a broad range of medical conditions, including infectious diseases such as influenza-like infections and acute viral hepatitis; cancer; endocrine, nutritional, and metabolic diseases; and diseases of the nervous, circulatory, digestive, respiratory, and musculoskeletal system. Examples of nervous system disorders associated with taste disorders include Alzheimer's disease, Bell's palsy, damage to chorda tympani, epilepsy, head trauma, Korsakoff's syndrome, multiple sclerosis, Parkinson's disease, and tumors. Examples of nutritional disorders linked with taste disorders include chronic renal failure, liver disease, including cirrhosis, niacin (vitamin B_3) deficiency, vitamin B_{12} deficiency, and zinc deficiency. Endocrine disorders associated with taste aberrations include adrenal cortical insufficiency, congenital adrenal hyperplasia, panhypopituitarism, Cushing's syndrome, diabetes mellitus, hypothyroidism, Kallmann's syndrome, pseudohypoparathyroidism, and Turner's syndrome. Cancer is an example of a medical condition in which patients are particularly vulnerable to taste disorders. Taste aberrations occur both in untreated cancer patients as well as those treated with radiation therapy or chemotherapy. Both malnutrition and damage to taste cell turnover (taste receptors replicate themselves every 10.5 days) by radiation and chemotherapy play a role in these taste changes in cancer patients. Some taste complaints by cancer patients, such as food aversions, are not due to altered sensory physiology per se but are a consequence of learned aversions in which the taste of foods is associated with the noxious effects of radiation and chemotherapy.

Causes

Drugs are the most common cause of bilateral taste dysfunction, and several studies suggest that up to 33% of elderly individuals experience medication-related alterations in taste. Hundreds of medications, including all major therapeutic drug classes, have been clinically reported to cause taste disorders in a percentage of people for whom they are prescribed. Typical taste complaints include "loss of taste," "altered taste function," and "metallic taste." It is not surprising that taste disorders are sequelae of medication usage because the function of drug therapy is to modify the biochemistry of cells to treat a diseased state or metabolic condition. Orally administered drug formulations can induce a taste of their own or modify transduction mechanisms in the taste receptor cells in the oral cavity or back of the throat. For some drugs, the plasma concentrations are sufficiently high to stimulate taste receptors from the basolateral side (near the blood) of taste cells (called intravascular taste). An example of a drug that can induce taste disorders from the basolateral side of taste cells is the protease inhibitor saquinavir that is used to treat AIDS. After a 600 milligram (mg) dose, the concentration of saquinavir in plasma is 0.22 millimolar (mM), which exceeds the lingual taste detection threshold of saquinavir of 0.0029 mM. Injectable drugs introduced into the blood circulation can also induce transient complaints of dysgeusia (e.g., metallic taste) at concentrations below taste threshold, and the taste complaints are often accompanied by nausea. The cause of dysgeusia in this case is due to activation (by the force of injection) of pressure receptors in the arterial system that are innervated by the IXth cranial nerve that projects to the nucleus of the solitary tract in the brain stem that also receives taste input. Even when salivary or plasma concentrations of drugs are lower than the threshold levels for taste perception, drugs and their metabolites can accumulate in taste buds over time to achieve concentrations that exceed taste thresholds and hence induce taste disorders. Drugs most likely to accumulate in taste buds are ones with high lipophilicity (fat solubility) that allows them to penetrate cell membranes and linger in taste cells. Examples include drugs such as amiodarone (antiarrhythmic), terfenadine (antihistamine), thioridazine (antipsychotic), atorvastatin (anticholesteremic), and pimozide (antipsychotic) for which the lipid solubility is millions of times greater than water solubility. Due to its high lipid solubility, the half-life of amiodarone in the body is 47 days or longer, which explains the observation that taste alterations persist for a long time after intake of the drug has ceased. Furthermore, high fat meals can increase the bioavailability of lipophilic drugs, such as the antiarrhythmic drug amiodarone, that is, the area under the plasma concentration time curve (AUC) and peak plasma concentration (C_{max}) are both increased when lipophilic drugs are consumed with a fatty meal. Elevated drug absorption increases the propensity for taste disorders.

A highly significant number of taste disorders occur as a result of drug–drug interactions as well as food–drug interactions that increase drug absorption above the recommended dose. The incidence of taste disorders increases with polypharmacy because it elevates the probability of these interactions. In order to understand how supratherapeutic concentrations of drugs result from polypharmacy, it is necessary to briefly describe the process called *first pass metabolism*. After a drug is swallowed, it is metabolized (biochemically modified or degraded) through specialized enzymatic processes located in the intestines and liver. The degree to which a drug is metabolized determines the bioavailability of the drug (i.e., the amount of the drug that reaches the circulatory system or target tissues). If a drug is co-administered with a second drug that inhibits the first drug's metabolism, this will decrease clearance of the first drug from the body. An example of this type of drug–drug interaction involves the anti-anxiety drug midazolam. When midazolam is co-administered with certain AIDS drugs (e.g., the protease inhibitor indinavir), the serum concentrations of midazolam (as well as the incidence of taste disorders) increase. This is due to the fact that AIDS drugs, such as indinavir, inhibit the metabolism and clearance of midazolam, allowing it to accumulate in the body and produce increased pharmacological responses. Drug–drug interactions (as well as taste disorders) can even occur after discontinuation of highly lipophilic drugs with a long half-life in the body. Genetic differences in genes for drug metabolism also account for the individual variability in predisposition to adverse taste effects from medications, (i.e. poor metabolizers of drugs are more susceptible to taste disorders due to reduced clearance of drugs from the body).

Burning Mouth Syndrome

Burning mouth syndrome (BMS), also referred to as glossodynia or glossalgia, is often considered as a taste disorder because the majority of people with this condition have altered taste function particularly related to the chorda tympani nerve that subserves taste perception on the anterior 2/3 of the tongue. BMS is a painful condition that is characterized by a burning or scalding sensation on the tongue, lips, palate, or throughout the mouth. Although BMS most commonly occurs in middle-age (especially menopausal) or older women, its etiology is unknown. Factors such as dry mouth (often caused by anticholinergic medications such as tricyclic antidepressants and diuretics), oral fungal infections (e.g., candidiasis), nutritional deficiencies, dental problems (including poorly fitting dentures), peripheral neuropathies, gastroesophageal reflux (GERD), angiotensin-converting enzyme (ACE) inhibitors, and medical conditions such as hypothyroidism and diabetes have been reported to precipitate the symptoms. Although it is not known if the burning symptoms are caused by a central or peripheral mechanism, current thought is that damage to the taste nerves diminishes the inhibition of trigeminal-nerve pain fibers, and this disinhibition in turn leads to oral burning symptoms.

Recovery and Treatment

The prognosis for recovery from taste disorders depends on the type of loss. Taste abnormalities experienced by older individuals in the absence of disease or medications are generally irreversible. Medication-induced taste alterations tend to be reversed with discontinuance of the drug, although resolution may take up to four months. For some drugs, however, long-term damage to taste perception persists (e.g., terbinafine, a common treatment for fungal infections of the toenail and fingernail) long after termination of drug therapy. Similarly, taste disorders subsequent to viral infections generally resolve, but some people experience permanent loss. Taste disorders can have a deleterious effect on nutritional status because taste cues play a major role in whether foods and beverages are swallowed. Furthermore, because taste signals from food elicit salivary, gastric acid, and pancreatic secretions, taste aberrations can blunt these digestive secretions and ultimately the absorption of nutrients. Whatever the cause, taste disorders can reduce appetite and the motivation to eat, interfere with the ability to modulate food intake as nutritional requirements vary over time, and impair quality of life.

Currently there are no effective treatments for persistent impairment of the sense of taste. In the future, use of stem cells that are capable of differentiating into taste receptors may become a therapeutic possibility. Patients with taste disorders sometimes find temporary relief from sipping water, sucking ice chips, avoiding spicy and acid foods, chewing sugarless gum, and/or accentuating other sensory stimuli, such as texture, temperature, and odor, at a meal.

Susan S. Schiffman

See also Taste: Genetics of; Taste: Supertasters; Taste and Food Preferences; Taste Receptors and Transduction; Taste Stimuli: Chemical and Food; Taste System Structure; Taste Thresholds and Intensity

Further Readings

Fukunaga, A., Uematsu, H., & Sugimoto, K. (2005). Influences of aging on taste perception and oral somatic sensation. *Journals of Gerontology Series A-Biological Sciences & Medical Sciences, 60*(1), 109–113.

Henkin, R. I. (1994). Drug-induced taste and smell disorders. Incidence, mechanisms and management related primarily to treatment of sensory receptor dysfunction. *Drug Safety, 11*(5), 318–377.

Schiffman, S. S. (1997). Taste and smell losses in normal aging and disease. *Journal of the American Medical Association, 278,* 1357–1362.

Schiffman, S. S., & Zervakis, J. (2002). Taste and smell perception in the elderly: Effect of medications and disease. *Advances in Food and Nutrition Research, 44,* 248–346.

TASTE: GENETICS OF

When any two people taste a food, even simple ones, they will not have the same sensory experience. And the perceptual differences between them will only grow with the complexity of the food they sample, such as with a full course meal served with wine. Biologists, philosophers, and psychologists have discussed this idea related to different species, stating that each must live in its own unique sensory world. And while the differences among humans may be more subtle than among species, the parallel seems clear that each human occupies a different perceptual and chemosensory world as well. Just as "color blind" people live in differently colored worlds than those who have full trichromatic color vision, so too do people live in different taste (and smell) worlds. But individual variations in taste and color perception differ in the numbers of people they impact. Color blindness affects a relatively small portion of the population, approximately 8% of men and 1% of women, whereas virtually all people live in their own taste and smell worlds. The determination of the degree to which these individuals' perceptual differences in taste can be attributed to differences in their genomes is the central focus of taste genetics.

There are many other factors that account for these individual differences in addition to genetics, including nutritional status, pre- and postnatal developmental history, hormonal status, time of day, as well as health, present drug treatment, and medical history. But one of the most common and potent determinants of individual taste differences are differences in people's genes. The protein receptors in our taste receptor cells respond to a wide variety of chemicals that are transduced into signals interpreted by our brains as sweet, salty, bitter, sour, and savory (umami) tastes. The genes that code for these taste receptors are highly variable and likely play a direct role in our individualized perceptions.

The study of taste genetics is important for several reasons. First, variations in our genes can determine how well we taste certain compounds. Second, how well we taste the chemicals that we perceive as sweet, sour, salty, bitter, and savory can impact whether we like the taste of a certain foods that contain these compounds. Third, whether we like the taste of a food can, in turn, affect whether we will eat it and, perhaps, how much of it; and this determines our nutritional status, which impacts our health, longevity, and quality of life. Importantly, the same sets of genes that are expressed in our mouths and affect our abilities to taste also code for proteins that are expressed throughout our gastrointestinal tract and can alter our "visceral taste." We do not consciously perceive visceral taste inputs, but the stimulation of these taste receptors by ingested foods alters hormone secretions and our metabolism, and can also have a large impact on our nutritional status.

When examining the taste responses to a single compound, the distributions of taste sensitivities can be narrow for compounds, such as NaCl (salty) and sucrose (sweet), or quite broad for many compounds, such as some bitter compound for which sensitivities are spread over 1,000- or 10,000-fold differences in concentration. The most well known and studied of these bitter compounds is the antithyroid drug phenylthiocarbamide (PTC). This entry discusses the taste receptors of the five major taste sensations: bitter, sweet, savory, sour, and salty.

Bitter Taste: The Phenylthiocarbamide (PTC) Story

Taste sensitivity to PTC has been known to vary among individuals since 1921. The differences in sensitivity are large, with some individuals tasting saturated solutions as only weakly bitter, whereas others perceive millimolar amounts as intensely bitter. This trait was quickly found to be heritable and has been studied in hundreds of published papers and among populations around the world, ranging from the Amhara of Africa to the Zagorsk of Russia. Today, small pieces of paper soaked in

PTC solutions can easily be purchased for classroom demonstrations of the extreme differences in students' abilities to taste the compound as bitter. PTC sensitivity is one of the oldest and most frequently studied perceptual traits in the world. It was studied in a great many different populations because tests were easy to administer and transport, and the test results were immediate and clear. Therefore, it was used as a genetic marker by anthropologists to characterize their study populations. But the variation in the genome that was responsible for this inherited taste ability eluded scientists until recently.

Approximately five years ago it was discovered that a member of the class of bitter taste receptors known as the class 2 taste receptors or TAS2Rs was responsible for the perceptual variation of PTC. This class of taste receptor has 25 members and is comprised of the membrane-bound receptors called 7-transmembrane spanning receptors (meaning the protein crosses the entire span of the membrane seven times). This receptor class is also coupled to G proteins, an initial amplifying component of the activated receptor's transduction sequence, giving rise to its common name of the G protein coupled receptors (GPCRs). The class of GPCRs that comprise the TAS2R bitter receptors are also characterized by their short amino terminal segment at the extracellular end of the protein. The TAS1Rs, another class of taste receptors, are also GPCRs, but they possess a very large segment at their extracellular amino terminal. All of these receptors display a large variety of forms within the population, which may well contribute to the observed variations in taste.

Genetics

The TAS2R that is responsible for perceptual variation of PTC is labeled *TAS2R38*. There are two very common forms of this gene worldwide. One form, or allele, of this gene is known as the PAV allele of TAS2R38 after the *p*roline, *a*lanine, and *v*aline amino acids at positions 49, 262, 296 of the receptor protein. The other allele is labeled AVI after the *a*lanine, *v*aline, and *i*soleucine at the same positions respectively. The PAV allele is associated with high sensitivity and strong responses to the compound PTC, and the AVI allele is associated with very weak responses to PTC. People who possess one PAV allele and one AVI allele display intermediate responses.

The different forms of the receptor that these alleles encode have also been expressed in a cell line maintained in a dish and have been functionally characterized for their responses to PTC. The receptors expressed in these cells showed sensitivities to PTC that closely paralleled the perceptual sensitivities of humans with two copies (were homozygous) of either the PAV or the AVI allele. Therefore, the gene that encodes these different receptor proteins is directly responsible for the large variations in perceptual sensitivity to the compound PTC.

Because all of these experiments were conducted with solutions containing the artificial compound PTC, the question arose whether these genetic differences would impact the perception of structurally related naturally occurring compounds in foods and vegetables. To answer this question, subjects were first screened to determine if they possessed the PAV-sensitive form of the gene or the AVI-insensitive form of the gene. These subjects were then given a wide variety of bitter tasting vegetables to taste and rate. Bitter vegetables either contained antithyroid compounds that are structurally related to PTC, such as those that occur in cruciferous vegetables like broccoli and Brussels sprouts, or possessed only structurally unrelated compounds, such as those that occur in bitter melon. While all subjects found all of the vegetables to be bitter, the two sets of subjects differed in their bitterness ratings of only the vegetables that contained the antithyroid compounds. Specifically, the subjects who were homozygous for the PAV allele found only those vegetables containing antithyroid compounds as more bitter than those who were homozygous for the AVI allele. Thus, subjects' differences in their *TAS2R38* receptor alleles selectively determined their perception of vegetables that contained compounds that activate this receptor.

Functional Significance

Why would people differ so profoundly in their taste perception of chemicals and foods, and why would this be encoded genetically? First, all humans originated in Africa and occupied a vast diversity of ecological niches there from coastal to mountainous and from jungle to savannah to desert. The variety of plants encountered across these different environments would be large, and among people who forage there would be a strong pressure to evolve mechanisms to detect poisons in plants. Second, as humans migrated out of Africa and spread around the world, the diversity of plants encountered was even greater, and the need to develop and diversify detector systems to avoid toxins was maintained and specialized.

So why would a population maintain different forms of a toxin detector that works well, as well as forms of the detector that do not work well? The answer here is less clear but may lie in the fact that what is toxic to some people is nutritious to others. Because the chemicals that activate the TAS2R38 receptor are largely antithyroid compounds, people who live within regions of endemic goiter and hypothyroidism and eat these toxins may severely damage themselves, especially developing children who would fail to develop normally both sexually and mentally. However, people who live near coastal areas who eat high iodine diets from seafood would not have to worry about hypothyroidism and could enjoy the vegetables that contain these compounds as highly nutritious.

There is an important piece of this puzzle still missing, however. It is clear that variations in taste receptor genes can affect our perception of certain taste compounds that act on the receptors coded by these genes. It is also clear that these genetic variations can even affect how we perceive foods. What remains to be determined is that variations in these genes affect how much we prefer and consume foods. These nutritional differences in turn will affect our health and well-being. The direct links among all of these are in part obscured by cultural practices that override our innate taste preferences.

There are several parallel stories unfolding for other bitter taste receptors as well. All 25 TAS2R genes are highly variable and some of them have sets of ligands identified. Thus, it seems likely that in the near future stories closely resembling the TAS2R38-PTC story will unfold for other TAS2R genes.

Sweet and Savory

There are only three members of the class I taste receptors, the TAS1Rs, and they are believed to play a central role in sweet and savory tastes. These receptors appear to operate principally with partners forming heteromers of TAS1R1-TAS1R3, an amino acid receptor, and TAS1R2-TAS1R3, a sweetener receptor. All three of these receptors are variable within the population, but the roles they play in human sweet and savory taste are unclear. The TAS1R2-TAS1R3 receptor appears to be important for mouse sweet taste, as variations in this receptor affect the ability of mice to perceive sweeteners such as saccharin. Also, natural variation in the *TAS1R2* gene of cats has rendered them unable to express the TAS1R protein. Consequently, cats appear to have great difficulty tasting sugars or any other sweeteners. Given the crucial role that TAS1R2-TAS1R3 appears to play for sweet taste in mammals, it will undoubtedly be true that this receptor is important for sweet taste in humans. Presently, there is little direct genetic evidence that variations in these two genes account for variations in sweet taste perception of any sweetener.

There are also large magnitude variations in people's ability to taste umami from glutamate salts or other amino acids. Some people, approximately 5% of the population among Europeans, appear to be profoundly insensitive to umami taste from glutamate salts. Again there are many variations in the human *TAS1R1* and *TAS1R3* genes, but these do not at this time account for this insensitivity to umami taste. Other taste receptors are thought to be involved in umami taste, such as the brain glutamate receptors mGluR1 and mGluR4. Their involvement in human umami taste remains undetermined.

Ionic Tastes: Sour and Salt

The two ions H^+ and Na^+ are almost solely responsible for sour taste and salty taste, respectively. Whereas other ions are perceived as salty, none are as clearly or purely salty as sodium, and there are few other stimuli besides protons that taste sour. Because the stimuli are ions, the presumed receptors are thought to be ion channels in taste receptor cell membranes. Sour taste is likely transduced by one or perhaps several proton channels in epithelial tissue. Salty taste is hypothesized to be transduced by epithelial sodium channels. However, these receptors have not yet been identified in humans as contributing to taste, and there is a distinct possibility that salty and sour receptors are not ion channels.

Both salty taste and sour taste in humans show rather wide variations in sensitivities in the population. Studies comparing identical twins, who share almost all of their genomes in common, with nonidentical twins, who share on average half of their genomes, reveal the sources of these two sets of variation to be quite different. The variation in sour taste sensitivity from acids is highly heritable, with identical twins perceiving acids more similarly than do nonidentical twins. The same analysis on salt taste variation shows that identical twins are not much more similar in their salty taste sensitivity than are nonidentical twins. Thus, the variation seen in sour taste sensitivity is largely genetically determined, whereas the variation seen in salty taste sensitivity is largely environmentally determined. Nevertheless, variations in both taste sensitivities display effects of environment and genes.

To date, physiologists and molecular biologists have elucidated much of the basic mechanisms of taste transduction in taste receptor cells that lead to taste perception. But for humans, as well as among many different species, taste perception varies considerably. In the simplest case, a single concentration of a single compound may be given to 100 random individuals and their responses to this taste solution will vary everywhere from "it tastes like water" to "it is intolerably strong" and a continuum of responses in between. Human genetic research of taste hopes to identify and understand the variants in the human DNA sequence that contribute to variations in taste perception. The beauty of the taste system as a subject of genetic research is that variation among individuals is ubiquitous and high in magnitude and that, in contrast to the vast majority of genetic effects, there are very large single gene effects that remain undiscovered. Ultimately, the nutritional and metabolic health consequences of this field of inquiry promise to be large and beneficial.

Paul A. S. Breslin

See also Individual Differences in Perception; Taste; Taste: Disorders; Taste: Supertasters; Taste and Food Preferences; Taste Receptors and Transduction

Further Readings

Breslin, P. A., & Huang, L. (2006). Human taste: Peripheral anatomy, taste transduction, and coding. *Advances in Oto-Rhino-Laryngology, 63,* 152–190.

Bufe, B., Breslin, P. A., Kuhn, C., Reed, D. R., Tharp, C. D., Slack, J. P., et al. (2005). The molecular basis of individual differences in phenylthiocarbamide and propylthiouracil bitterness perception. *Current Biology, 15,* 322–327.

Deshpande, O., Batzoglou, S., Feldman, M. W., & Cavalli-Sforza, L. L. (2009). A serial founder effect model for human settlement out of Africa. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 276,* 291–300.

Drayna, D. (2005). Founder mutations. *Scientific American, 293,* 78–85.

Guo, S. W., & Reed, D. R. (2001). The genetics of phenylthiocarbamide perception. *Annals of Human Biology, 28,* 111–142.

Hansen, J. L., Reed, D. R., Wright, M. J., Martin, N. G., & Breslin, P. A. (2006). Heritability and genetic covariation of sensitivity to PROP, SOA, Quinine Hcl, and caffeine. *Chemical Senses, 31,* 403–413.

Kim, U. K., Jorgenson, E., Coon, H., Leppert, M., Risch, N., & Drayna, D. (2003). Positional cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. *Science, 299,* 1221–1225.

Sandell, M. A., & Breslin, P. A. (2006). Variability in a taste-receptor gene determines whether we taste toxins in food. *Current Biology, 16,* R792–R794.

Wang, X., Thomas, S. D., & Zhang, J. (2004). Relaxation of selective constraint and loss of function in the evolution of human bitter taste receptor genes. *Human Molecular Genetics, 13,* 2671–2678.

Wise, P. M., Hansen, J. L., Reed, D. R., & Breslin, P. A. (2007). Twin study of the heritability of recognition thresholds for sour and salty taste. *Chemical Senses, 32,* 749–754.

Wynne, C. D. L. (2004). *Do animals think?* Princeton, NJ: Princeton University Press.

TASTE: SUPERTASTERS

Although there are individual differences in the extent to which we are sensitive to all taste qualities, such variation is most pronounced for bitterness. Human responses to bitter thiourea compounds, in particular *phenylthiocarbamide* (PTC) and

6-n-propylthiouracil (PROP), have been extensively studied for several decades, and it is well established that 25 to 30% of Caucasian populations are highly insensitive to their taste (*nontasters*), although this proportion does vary across cultures. Among those who find PTC/PROP bitter (*tasters*), there is also considerable variation in both thresholds and suprathreshold ratings of intensity. In 1992, Linda Bartoshuk and colleagues at Yale University reported that a subgroup of tasters (estimated also as 20–25% of the population, or around a third of all tasters) rated PROP as profoundly bitter. Such *supertasters* may have been previously overlooked, due to the initial widespread use of detection (threshold) criteria for PTC/PROP taster group classification, which gave no indication of variation in suprathreshold responses. This entry discusses the genetic variations, the perception of somatosensory qualities, and food preferences and health implications as they relate to supertasters.

Genetic Variations

The existence of a very wide range of chemical structures that are bitter to humans has suggested the possibility of many different forms of receptor, and this has largely been confirmed by the discovery of the *TAS2R* family of bitter receptor genes. Recent studies of taste genetics have identified specific polymorphisms (structural variations) in the bitter taste receptor gene TAS2R38 on chromosome 7 that are associated with specific PTC/PROP taster/nontaster phenotypes. Those individuals who are heterozygous for the taster genotype are also tasters, although less sensitive than homozygous tasters. Although women are more sensitive to PROP than are men and are more likely to be supertasters, the inheritance pattern does not appear to be sex-linked.

Genetic variations in responses to PROP (now much more widely studied than PTC) would be of limited interest from the point of view of understanding taste perception if it were not for the fact that the different PROP taster groups—nontasters, medium-tasters, and supertasters—also vary considerably in response to a wide range of other tastes, including not only those of other bitter compounds, but sweet, salty, and sour tastes as well. In studies of these relationships to date, supertasters are found to respond to tastes—both alone and in mixtures, as well as in foods and beverages—with both significantly high ratings of intensity and greater sensitivity to intensity variations, than are either medium-tasters or nontasters. The higher responsiveness of supertasters to tastes in general also influences taste mixture interactions, shown, for example, in more pronounced suppression of sweet tastes by bitterness than occurs in other taster groups.

Perception of Somatosensory Qualities

Strong relationships between PROP tasting and the perception of oral somatosensory qualities have also been established. Sensations produced by oral irritants, such as capsaicin (the primary irritant in chilies), carbonation in drinks, acids, and ethanol, as well as the tactile sensations associated with viscosity or fat content in foods and beverages, show a positive relationship with ratings of PROP intensity. Supertasters tend to show the highest ratings of these qualities and greatest sensitivity to variations in fat content and viscosity within foods.

The relationships between PROP tasting and the perception of other taste and somatosensory qualities appear to be mediated via an association with individual variations in taste anatomy. The density of fungiform papillae (the structures that contain the collections of taste cells as *taste buds*) on the anterior tongue can vary across individuals by as much as 16-fold, and there are significant positive associations between PROP intensity ratings and density of both fungiform papillae and taste pores (through which tastants reach the receptors). A greater than 5-fold increase in taste pore density between nontasters and supertasters has been reported. Because taste intensities reflect spatial summation—that is, they are at least partly a function of the number of fungiform papillae and taste pores stimulated—the associations between PROP tasting and fungiform papillae density may be sufficient to explain the relationships seen between PROP tasting and suprathreshold measures of the intensity of other tastes. They also account for the relationship with oral somatosensory qualities because fungiform papillae are heavily innervated not only by taste afferents within the chorda tympani (VIIth) nerve, but also by trigeminal (Vth) nerve fibers mediating chemesthetic (oral irritation) and tactile qualities. Hence, greater fungiform papillae density will mean higher densities of fibers responsive to oral irritation produced by capsaicin or alcohol, or tactile qualities such as viscosity. The magnitude of the differences in fungiform papillae densities between taster groups may mean that they might be a better explanation of differences in ratings of both tastes in general and oral somatosensory qualities than are ratings of PROP taste.

Both molecular and psychophysical evidence suggests that neither variations in the TAS2R38 gene nor fungiform papillae density alone can explain the uniquely broad variation in measures of PROP intensity. When grouped according to genotype, tasters show considerable variation in their responses to PROP/PTC, in contrast to nontasters who tend to be more tightly grouped. Moreover, differences between homozygous and heterozygous taster genotypes in suprathreshold ratings of PROP are only modest, in contrast to the often very much higher ratings of bitterness shown by supertasters rather than medium-tasters.

In fact, the relationship between fungiform papillae density and PROP intensity does not appear to hold for nontasters, whereas there *is* a relationship with quinine bitterness. This suggests strongly that PROP nontasting is primarily accounted for by purely genetic factors. By contrast, PROP tasters show positive correlations between fungiform papillae counts, and both quinine and PROP intensity. Hence, for the majority of bitter compounds (and likely for other tastes as well) spatial summation over distinctly varying fungiform papillae densities is likely to account for variation in intensity among tasters. Supertasters may therefore be those tasters who are homozygous for the taster allele *and* who also have a large number of fungiform papillae. Conversely, there are individuals who are supertasters in terms of their responses to most tastes and irritant qualities, but are PROP nontasters. In effect, then, the relationships seen between PROP and other taste intensities are based on fungiform papillae numbers, rather than PROP sensitivity per se. This conclusion is supported by findings that receptors expressing the TAS2R38 taster haplotype are unresponsive to other bitter compounds.

Food Preferences and Health Implications

PROP taster differences in psychophysical responses to sensory qualities have a wider significance in that they are also reflected in food preferences. Tasters have more food aversions than nontasters, and because of their responses to prominent tastes—particularly bitterness and sourness—or oral irritants, the prototypical supertaster tends to dislike, for example, black coffee, hot (spicy) foods, grapefruit juice, and other bitter fruits and vegetables. This has led to interest in the PROP phenotype as a predictor of food choices that have known health implications. In particular, tasting vegetables as more bitter is associated with both reduced preference and intake. This is important in the case of bitter green, leafy vegetables, which contain high levels of antioxidant compounds thought to be protective against certain cancers. Thus, it has been suggested that supertasters, by virtue of their greater response to bitterness in foods, are relatively more susceptible to disease states for which food choice can be a risk factor. This is supported by reports of increased counts of precancerous colon polyps as a function of higher PROP ratings. Conversely, supertasters may be "protected" when it comes to excessive alcohol use, due to their perception of greater bitterness and irritation (burn) from many alcoholic beverages, leading to lower intake. Nevertheless, evidence for connections between taste genotype/phenotype and disease, and for the inverse relationships between PROP tasting and body mass index that have been reported, is still somewhat limited. However, such relationships—if confirmed—are strong indicators that individual differences in taste perceptions can be an important influence on not just what we prefer, but also of the health consequences of our food choices.

J. Prescott

See also Individual Differences in Perception; Taste; Taste: Disorders; Taste and Food Preferences; Taste Receptors and Transduction

Further Readings

Bartoshuk, L. M., Duffy, V. B., Reed, D., & Williams, A. (1996). Supertasting, earaches and head injury: Genetics and pathology alter our taste worlds. *Neuroscience & Biobehavioral Review, 20*(1), 79–87.

Bufe, B., Breslin, P. A. S., Kuhn, C., Reed, D. R., Tharp, C. D., Slack, J. P., et al. (2005). The molecular basis of individual differences in phenylthiocarbamide and propylthiouracil bitterness perception. *Current Biology, 15,* 322–327.

Hayes, J. E., Bartoshuk, L. M., Kidd, J. R., & Duffy, V. B. (2008). Supertasting and PROP bitterness depends on more than TAS2R38 gene. *Chemical Senses, 33,* 255–265.

Prescott, J., & Tepper, B. (Eds.). (2004). *Genetic variation in taste sensitivity.* New York: Marcel Dekker.

TASTE ADAPTATION

Although the saliva that constantly bathes our tongue contains a number of taste stimuli, we ordinarily find it quite tasteless. However, if you were to rinse your tongue with water for a minute or so, you would find your own saliva to have a distinctly salty taste. And if you rinsed the tongue with a salt solution only somewhat saltier than your saliva, pure water would then taste distinctly sour-bitter. Thus, the tongue is constantly adapts to the stimuli that surround it and responds to changes from the adapting level. The range over which adaptation is complete is at least two orders of magnitude, or more than 100 to 1.

As in all the senses, adaptation in taste has many important functions. Most obvious is the role of adaptation in tuning out steady stimulation, as in the previous example. But this is only one of a number of functions of adaptation found in the senses in general. The effects discussed here tend to have their analogs in the other senses.

After adapting to a given concentration of salt, the tongue is better able to detect changes in concentration near the adapting level than it would have been without adaptation. In other words, the difference threshold is smaller near the adapting concentration. In addition, the tongue responds more to increases in concentration than it did before adaptation. Thus, the rate of increase in sensation with increasing concentration becomes steeper. So after adaptation, the threshold concentration required to produce a sensation is greater than before, but the sensations caused by stronger concentrations are less affected.

These two effects illustrate how adaptation acts as a sort of range-finding mechanism that adjusts its sensitivity up or down in order to keep the system as responsive as possible. This is similar to the way a camera adjusts to changing light levels. Thus, it is misleading to think of adaptation primarily as a loss of sensation. The sensation caused by a given stimulus does decrease with adaptation, but the system is *more* sensitive to *changes* in stimulation near the adapting concentration than it was before adaptation. This entry describes adaptation as temporal sensitivity, water taste, adaptation and the basic tastes, importance of method of stimulus presentation, and the relation between adaptation and habituation.

Adaptation as Temporal Sensitivity

Like all biological processes, sensations take place in time. Thus, it can make a considerable difference at what point during the presentation of a stimulus that a response is measured. Taste sensation takes several seconds to grow to a maximum before it begins to decline. The rate of adaptation is similar for most compounds when starting subjective intensities are the same, and adaptation is complete after about two minutes. Not only is the time to complete adaptation longer for stronger concentrations, but the rate of decline is also slower. In other words, the time constant of the exponential function describing the function is longer for stronger concentrations. For artificial sweeteners, the time course of the response is slower, both in the rise and decline of intensity.

The most obvious consequence of this time dependency is that there is general recognition among taste researchers that it is necessary to rinse the tongue to remove the stimulus and wait for it to recover from its effects before presenting another stimulus. But the taste of a second taste stimulus following a previous one will be the result of interactions between the declining effects of the first stimulus and the growing effects of the second if they follow each other closely, even if the tongue has been rinsed.

Water Taste

Water is only tasteless when the tongue is adapted to water, an unusual event. When the tongue is adapted to various taste compounds, water can take on any of the four traditional qualities: salt, sour, sweet, and bitter. Adaptation to salt causes water to have a sour-bitter taste; both sour and bitter adaptation leads water to taste sweet; and sweet adaptation causes water to taste bitter. Only adaptation to urea causes water to taste salty. It would be tempting to see this as parallel to the phenomena seen with color afterimages.

In the case of color, the interactions between different qualities are symmetrical: If one presents green, one gets an aftereffect called an afterimage (like the floating spots that occur after looking at a camera's flash) that looks red. If red is presented, then a green afterimage follows. Similarly, blue causes a yellow afterimage; and yellow causes a blue afterimage. These visual effects have been explained in terms of mechanisms in which these pairs of colors are linked together. However, the situation is different in taste. Although there are some symmetrical adaptation effects (sweet adaptation causes water to taste bitter, and bitter adaptation causes water to taste sweet), there are other situations in which the effects are not symmetrical (adaptation to salt causes water to taste sour-bitter, whereas adaptation to sour-bitter substances does not cause water to taste salty). What these nonsymmetrical effects suggest is that the mechanisms between different taste qualities are not linked in the same way as are the mechanisms for some colors. Additional evidence suggesting that mechanisms responsible for different tastes function independently is provided by experiments that involve the four basic tastes.

Adaptation and the Basic Tastes

This section considers adaptation to salty, sour, sweet, and bitter, the traditional basic tastes. (Adaptation to monosodium glutamate [MSG], which gives reports of *umami,* the proposed fifth basic taste, has not been studied.) Cross-adaptation is tested when a different stimulus is presented after adaptation to a previous stimulus. For example, one can present NaCl, which has a pure salty taste, and then present other salts that have other (mostly bitter) side tastes as well as tasting salty. If adaptation to one compound reduces the taste of another that follows, then this is taken as evidence that the two compounds are subserved by the same physiological mechanism.

Additionally, cross-adaptation has been studied across taste qualities. For example, one can adapt the tongue to a stimulus such as NaCl and test stimuli, such as sugar, that have other tastes. Cross-adaptation is not found between stimuli having different tastes, implying that the mechanisms subserving different taste qualities are separate. But adaptation to one stimulus will add the water taste of that stimulus to a following stimulus. Thus, adaptation to NaCl, a pure salty compound, will cause other salts to lose their salty taste; but the other salt will also be more bitter and sour than before adaptation because of the water taste that follows salt adaptation. Similar interactions occur after adaptation to sweet, bitter, and sour compounds.

Studies of adaptation of taste have been complicated by the fact that many compounds have side tastes. NaCl is almost the only pure salty substance, and most other salts also taste bitter. Most sweet substances other than sucrose also have bitter side tastes. It is necessary to measure magnitude of the various taste qualities, not just overall magnitude, in order to determine cross-adaptation effects. Similarly, measuring the threshold of one substance after adaptation to another substance needs to take quality into account.

Adaptation to a stimulus representing one of the basic tastes tends to reduce that taste in other stimuli. For example, adaptation to NaCl, which is the prototypical salty stimulus, reduces the apparent saltiness of all other salts. The bitter side tastes of those salts remain. Adaptation to citric acid reduces the sourness of all other sour stimuli. The situation is more complex for sweet and bitter substances. Adaptation to sucrose or other sugars reduces the intensity of other sweet substances, but when artificial sweeteners are the adapting compound, cross-adaptation may reduce the intensity of the test stimulus only partially.

Cross-adaptation between bitter compounds is also not always complete; this is consistent with the diverse nature of bitter taste stimuli. Physiological and genetic studies reveal a number of different cellular mechanisms that subserve taste responses: at least 25 for bitterness alone. The fact that psychophysical studies find cross-adaptation across stimuli that are coded by different cellular mechanisms implies that cross-adaptation occurs more centrally than the receptors.

As previously mentioned, the time course of adaptation for artificial sweeteners is longer than for natural sweeteners; this suggests that the coding mechanisms for the two types of sweeteners are different. It is also possible that they all stimulate the same sweet mechanism, but the artificial sweeteners additionally stimulate another mechanism. It should be noted that similar time courses for representatives of the other taste qualities do not rule out different coding mechanisms.

Importance of Method of Stimulus Presentation

How the stimulus is presented affects the outcomes of adaptation. In the usual method of sipping a stimulus, holding it in the mouth, and then spitting it out (commonly known as "sip and spit"), the stimulus spreads uncontrolled over parts of the tongue and mouth. The stimulus also mixes with saliva, which dilutes the stimulus as well as adding other potential tastants. This method has the advantage of simplicity and naturalness, and is understandably common in food-industry inspired research. However, more control over the area stimulated and the prevention of saliva from mixing with the stimulus is obtained by flow systems, usually presenting stimuli to the tongue outstretched between closed lips. Flow-system studies find more complete adaptation and more clear effects on the taste intensity and quality of stimuli presented after such adaptation.

The Relation Between Adaptation and Habituation

Capsaicin is the active ingredient in chili peppers. Studies of the pain caused by capsaicin applied to the tongue have yielded insights into adaptation. Traditionally, the term *adaptation* has been used in the context of sensory processes. It is considered to be rapid, reversible, and accomplished by peripheral (receptor) mechanisms. *Habituation* is the term used in the context of learning for similar declines in response, such as the reduction in the startle response over repeated stimulation. It is thought to be slower, less reversible, and central in its locus (i.e., not a sensory phenomenon).

But the results of adaptation to capsaicin call the distinction between adaptation and habituation into question. When capsaicin is presented continuously, adaptation takes the better part of an hour. The burn first increases for about 15 minutes, and then takes about a half hour to decline (and some subjects do not show a decline at all during that time). After adaptation to capsaicin, the response to later presentations of capsaicin is reduced for at least a week. This helps to explain why some people can tolerate far hotter foods than others. But these facts make time course a problematic criterion of defining adaptation because such long-term effects are thought to be characteristic of habituation.

Donald H. McBurney

See also Selective Adaptation; Taste; Taste Stimuli: Chemical and Food

Further Readings

Bartoshuk, L. M., McBurney, D. H., & Pfaffmann, C. (1964). Taste of sodium chloride solutions after adaptation to sodium chloride: Implications for the "water taste." *Science, 143,* 967–968.

Gent, J. F. (1979). An exponential model for adaptation in taste. *Sensory Processes, 3,* 303–316.

Keast, R. S. J., & Breslin, P. A. S. (2002). Cross-adaptation and bitterness inhibition of L-tryptophan, L-phenylalanine and urea: Further support for shared peripheral physiology. *Chemical Senses, 27,* 123–131.

McBurney, D. H. (1972). Gustatory cross adaptation between sweet-tasting compounds. *Perception & Psychophysics, 11,* 225–227.

McBurney, D. H., & Bartoshuk, L. M. (1973). Interactions between stimuli with different taste qualities. *Physiology and Behavior, 10,* 1101–1106.

McBurney, D. H., & Pfaffmann, C. (1963). Gustatory adaptation to saliva and sodium chloride. *Journal of Experimental Psychology, 65,* 523–529.

Schiffman, S. S., Cahn, H., & Lindley, M. G. (1981). Multiple receptor sites mediate sweetness: Evidence from cross adaptation. *Pharmacology, Biochemistry, and Behavior, 15,* 377–388.

TASTE AND FOOD PREFERENCES

Tastes can change our feeding behavior, either by motivating further eating or by leading to rejection of a food or beverage. This capability of shifting behavior is accompanied by positive or negative affective responses to those foods and beverages: We like or dislike them. By origin a biologically relevant facility, taste perception has a great impact on the food industry and society. Food-product development is directed to produce tasty foods worldwide, to gain popularity among consumers, and thus to increase sales. At the same time, easy access to a wide supply of good-tasting foods creates a risk for human health and well-being, as such foods are likely to motivate excess consumption or bias choices toward unhealthy options.

Tastes are the core of sensory properties of foods. In common language, this is underscored by the dichotomous division of foods into "sweet" and "salty" (or savory) categories. Almost any food or beverage enables the perception of sweet, salty, sour, bitter, or umami taste, and tastes are often present in combinations. Sweet and sour, sometimes accompanied by bitter, are salient tastes in many fruits and berries. Many savory foods provide both salty and umami taste. Sweet and sour appear as an alliance in Chinese cuisine, and salty and sour tastes are combined in pickled foods.

A single taste or a combination of tastes are embedded in a product that often also provides a complex combination of aroma, texture, and appearance. These "symphonies" of sensory properties form the sensory character of a particular food that makes the food recognizable in a given food culture. Availability of ingredients, the cooking and preservation conditions, as well as the ways of setting up and serving meals vary by cultures, often geographically. Different cuisines result in different sensory profiles of common foods, whereby the taste intensities also vary. For example, in some regions, bread can be prepared without any sodium chloride, thus appearing without salty taste, while in other parts of the world it appears distinctly salty due to high sodium chloride contents. Bread can also taste sweet or sour, depending on regional bread-making practices around the world. This entry discusses inherent taste preferences, developmental and cultural influences on taste preferences, cognitive inputs to taste perception and preference, and testing consumer responses to tastes.

Inherent Taste Preferences

Newborn babies display a variety of responses to tastes. Sweet taste is favored from the beginning of life, as evidenced by facial reflexes, changes in sucking movements, and consumed volumes of sweetened relative to plain water. Perception of sour or bitter taste results in rejecting facial expressions. Salty taste does not induce a positive nor negative facial expression in a newborn, but a few months after the birth, a baby prefers salty over plain water. Umami taste induces a positive response in children less than two years old, but only in combination with other flavors. Such responses provide a basis on which taste preferences for foods and beverages will be founded; however, adaptive changes will soon appear. A huge amount of exposure to foods and enormous physical, emotional, and cognitive development within a person will take place as taste preferences are established and fixed to culturally available and thus, acceptable, foods.

Taste preferences in young children are assumed to be universal—everybody favors sweet and rejects bitter and sour. However, individual responses to taste vary in their strength, and recent research suggests that a genetic component underlies such differences. Hence, the presence or absence of a "sweet tooth" would steer individual responses to sweet products. Similarly, a genetically determined disposition would impact preferences and consumption of vegetables with a bitter note. Furthermore, the ability to taste a bitter-tasting medicinal compound PROP (6-*n*-propylthiouracil) has been suggested to predict taste perceptions and preferences.

Developmental and Cultural Influences on Taste Preferences

In light of the survival value, it is not surprising that sensory characteristics of foods and beverages are subject to early learning. Young infants have an enormous capacity to learn. Regarding food perception and preference, several learning mechanisms are well documented, but their relative contributions are less clear. Repeated exposures are the key for learning to like tastes. This was well demonstrated in a study in which two groups of preschool children were repeatedly exposed to either sweet or salty tofu. During several exposures, the initial preference for the unflavored tofu switched to higher liking for the sweet or salty tofu, depending on the exposed flavor. A learned preference for a taste in a certain product does not necessarily generalize to a preference for that taste in other contexts: Exposing babies to sweetened water led to liking for sweetened water later in life, but did not predict liking for sweetness in a beverage. Salty water, well accepted by babies at a few months, was rejected at the age of two to three years, whereas salty food was preferred at this age. These examples suggest that a child learns to like a food as an entity of which a taste is a part. Recent brain research supports these experimental findings: The brain responds to the stimulation of different sensory modalities (e.g., visual, smell, and taste) in a multimodal fashion so that a response to, for example, sweetness depends on textural properties of the carrier (the food). However, research also shows that exposures to sour-tasting baby formulas may lead to better acceptance of sour taste in other foods later in childhood. Present research cannot answer the question concerning the critical age or context leading to long-term generalized food preference.

Everyday life shows that bitter-tasting products, such as coffee, beer, and grapefruit, gain wide acceptance when people grow older. In general, humans are flexible and capable of overcoming their neophobic responses to foods and, thus, expanding their preferences. A basis for expanding preferences is repeated exposures, underpinned by social support (e.g., pressures from parents or peers) to accept an initially poorly liked food. The association of a less-liked taste with another, well-liked sensory property (such as sweetness), or favorable physiological consequences of ingestion (e.g., stimulation from caffeine in coffee or alcohol in beer) are helpful in shaping positive responses. Thus, learning to like bitterness most likely involves a variety of physiological, psychological, and social mechanisms.

Regarding the life span, there are indications that, in comparison to adults, children and adolescents favor sweetness. Thus, high sweetness is preferred during the physical growth and could be explained by the need for energy. Elderly people may also tend to prefer high sweetness. This is unrelated to changes in perceptual capacity, as taste perception, unlike odor, does not undergo major changes at older age.

Preferences for tastes vary by culture or subculture. The variations can be attributed to the exposures to foods and beverages of a certain quality. Cross-cultural comparisons of Australian and Japanese subjects suggest that, along with the quantity of taste substance and the consequent taste intensity, familiarity influences the overall liking of a food product. This emphasizes once again how, in the course of familiarization, a taste and a certain food form an inseparable blending in the perceptual world of a consumer.

Within a culture, individual variations are also present. As examples, habitual sweetening of coffee varies within cultures, and large variations in affective responses to sourness exist. Such segmentation may be driven by availability and exposures, but inherent, genetically driven preferences have not been excluded to date.

Cognitive Inputs to Taste Perception and Preference

For consumers, tastes have meanings beyond sensory perceptions. Health-related concerns about taste substances create conflicts and may have an impact on preferences and consumption of sweet, salty, and umami tasting foods.

Sugars (in particular sucrose, the most common sweetener in foods and beverages) do not have useful nutrients other than energy, thus they are a source of "empty calories." They are also associated with being overweight, although evidence on this issue is weak. Excessive use of sucrose, combined with poor dental hygiene, leads to dental caries, particularly in vulnerable groups, such as children and teenagers. Excessive use can also bias the nutritional quality of diets in these groups. Dietary guidelines recommend worldwide restrictions the use of sugars. Yet sweetness is a great source of pleasure to many people, and special sweet products, such as chocolate, can be a target of craving or addiction. Artificial sweeteners have been developed to resolve the conflict, and they provide the proper sweetness in many beverages reasonably well. They are far less successful in solid foods, which is at least partially due to the crucial contribution of sugar to food texture.

Although sodium is a necessary substance for the body, excessive sodium chloride consumption is common in industrial countries, and dietary guidelines consistently advise cutting the consumption. The major health risks associated with excess consumption of sodium are elevated blood pressure and cardiovascular diseases. Salts (other than sodium salts) are available in commercial mineral salts aimed to act as sodium substitutes, but no perfect substitute is available. The most viable approach to reducing sodium intake is to pay attention to the sodium content of commercial products, choose low-salt alternatives, and gradually learn to like less saltiness. Controlled long-term studies show that exposure to lowered sodium intake leads to lowered preferred saltiness in up to three months.

Umami is the characteristic taste of monosodium glutamate (MSG), sodium salt of glutamic acid (amino acid), although this is not the only source of umami taste. Umami taste enhances the overall flavor of meat, fish, and vegetable products, for which it is widely used as a food additive. However, MSG has a long history of evoking negative consumer reactions because of perceived adverse health effects. Safety reviews show that the adverse effects have little factual base. Nevertheless, manufacturers try to avoid the use of MSG when possible, and instead substitute mixtures of amino acids and their salts.

The conflict experienced by the consumer, when facing the pressures to decrease the consumption of salient taste substances, is understandable because the established food preferences are an outcome of a long-term process. The meanings attached to tastes, the habitual choice of certain foods and beverages, and the established expectations are all menaced by the pressure to change. If the pressure is acknowledged, strong, new alternatives are sought. Research on the relationship between consumption practices and affective responses to foods suggests that a change in preference will follow in a delayed fashion when actual choices have enabled exposures to lower sweetness or saltiness.

Testing Consumer Responses to Tastes

Because of the central position of tastes among sensory properties of foods, consumer responses to tastes are vital information for food producers who try to maximize the success of their products. Such information is also crucial for health and nutrition experts who want to understand barriers to changes in food consumption.

Testing consumer responses to tastes are basically a psychophysical measurement, as respondents evaluate stimuli—samples of foods or beverages. Yet consumer testing involves elements that set testing apart from "pure" psychophysical testing typically occurring in a research laboratory. These elements improve the external validity of the responses, important for the application of results. Yet they impose limited extrapolation on the outcome. Features characteristic of consumer studies are discussed next.

Stimuli

In laboratory-based psychophysical testing, the taste stimuli are typically aqueous solutions of taste substances. The resulting data are expected to be "universal," free of noise that follows from using a perceptually complex food system. By contrast, complex food systems serve as stimuli in consumer tests because responses to the entire product, containing a given taste substance, are of interest. When the food system becomes more complex—assumingly, the least complex system is water solution—the taste substance is perceived differently. For example, sweetness differences are considerably more difficult to perceive in orange juice than in water; and the use of carbonated, instead of plain water, decreases the perceived sweetness. Research shows that reducing sucrose in yogurt leads to higher perceived sourness and less thick or less fatty mouthfeel. Interactions are specific for each food or beverage system. Numerous and somewhat unpredictable interactions prevent generalization to other foods or beverages or over the category of sweet products.

Respondents

Demands for human subjects recruited to consumer tests differ from those recruited to typical psychophysical testing. Respondents are usually chosen based on a few characteristics that define them as representatives of their segment of consumers. Thus, soft drinks samples are presented to teenagers, not to elderly people who rarely consume them. Products intended for the elderly (e.g., those with specific health effects) are tested using elderly consumers. The inclusion criteria for respondents often comprise age, gender, and minimum consumption frequency of the tested product or category.

The number of consumer panelists needed is usually larger than a group of respondents needed in psychophysical testing. Consumer responses are dispersed and subject to random variation. The panel size makes up for this, making it possible to find a trend and draw a valid conclusion despite the noise. Roughly, the size of a consumer panel should be 30 to 50 respondents per subgroup.

Rating Tasks

Consumer testing focuses on affective responses to products. This is due to the purpose of testing, as the data are used for the prediction of product success. Product information not related to affection (e.g., perceived intensities or differences of food samples) is also needed, but such data are more typically collected from trained sensory panels in controlled laboratory conditions.

A background questionnaire, to be completed by respondents, may include individual ratings of liking and use frequencies of sweet or salty foods, to be used for the construction of instruments that quantify consumer preferences. Instruments tracking relevant attitudes and mental dispositions may also be useful. The instruments can shed further light to otherwise unexplainable variations in consumer responses, as tastes—and foods in which they are experienced—may evoke different responses, depending on food consumption practices, values, beliefs, attitudes, and expectations of an individual. This is particularly true for sweetness, saltiness, and possibly umami because they are subject to conflicting opinions.

Location

Testing may take place in the company, research institute, or another well-controlled facility. Testing in a busy environment, for example in a shopping mall where respondents are recruited in the middle of their shopping activities, is common, but increases the risk of dispersion in responses. Testing can also be conducted at home so that respondents use a food or beverage over a certain period and fill out questionnaires on paper or the Internet while testing. Home-use testing is more natural than a shopping mall test, but prone to biases and distractions due to lack of a controlled environment.

Despite the differences from psychophysical approach, consumer testing has a goal similar to any psychophysical testing: to enable sound conclusions based on useful and valid data.

Hely Tuorila

See also Taste; Taste: Genetics of; Taste: Supertasters; Taste Stimuli: Chemical and Food; Taste Thresholds and Intensity

Further Readings

Bellisle, F. (1999). Glutamate and the UMAMI taste: Sensory, metabolic, nutritional and behavioural considerations. A review of the literature published in the last 10 years. *Neuroscience and Biobehavioral Reviews, 23,* 423–438.

Keskitalo, K., Knaapila, A., Kallela, M., Palotie, A., Wessman, M., Sammalisto, et al. (2007). Sweet taste preferences are partly genetically determined; identification of a trait locus on chromosome 16. *American Journal of Clinical Nutrition, 86,* 55–63.

Kilcast, D., & Angus, F. (Eds.). (2007). *Reducing salt in foods: Practical strategies.* Cambridge, UK: Woodhead.

Liem, D. G., & Mennella, J. A. (2003). Heightened sour preferences during childhood. *Chemical Senses, 28,* 173–180.

MacFie, H. (2007). *Consumer-led food product development.* Cambridge, UK: Woodhead.

Prescott, J., Bell, G. A., Gillmore, R., Yoshida, M., Sullivan, M. O., Korac, S., et al. (1998). Cross-cultural comparisons of Japanese and Australian responses to manipulations of sourness, saltiness and bitterness of foods. *Food Quality and Preference, 9,* 53–66.

Rozin, P. (2006). The integration of biological, social, cultural and psychological influences on food choice. In R. Shepherd & M. Raats (Eds.), *The psychology of food choice* (pp. 19–39). Oxfordshire, UK: CABI.

Sandell, M. A., & Breslin, P. A. S. (2006). Variability in a taste-receptor gene determines whether we taste toxins in foods. *Current Biology, 16,* R792–R794.

Spillane, W. J. (Ed.). (2006). *Optimizing sweet taste in foods.* Cambridge, UK: Woodhead.

Sullivan, S. A., & Birch, L. L. (1990). Pass the sugar, pass the salt: Experience dictates preference. *Developmental Psychology, 26,* 546–551.

Tepper, B. J. (2008). Nutritional implications of genetic taste variation: The role of PROP sensitivity and other taste phenotypes. *Annual Review of Nutrition, 28,* 367–388.

TASTE RECEPTORS AND TRANSDUCTION

Eating serves a basic biological role—a key reason we eat is to acquire nutrients necessary for our health and survival. The sense of taste influences what we eat by helping us to judge a food's quality before we swallow it. Sweet and savory tastes suggest the presence of carbohydrates and proteins, important nutrient sources. Salty taste indicates that foods contain sodium chloride or other salts necessary for proper cellular function. Bitter taste is often the sign of toxins; sour taste may indicate spoilage. Of course, how we use this taste information depends on our experience, our preferences, and our motivations. For example, we may avoid sweet-tasting foods because we want to lose weight, and we can learn to enjoy bitter-tasting drinks like beer and coffee. Before we can make any of these judgments, we must first detect the taste stimuli present in the foods we eat. Humans and other mammals use a small number of receptor proteins, present on the surface of sensory cells in the oral cavity, to recognize taste stimuli (called *tastants*). Upon interaction with these tastants, the activated taste receptors trigger a cascade of molecular interactions within the sensory cell to transduce, or change, the chemical signal into a cellular signal that can be communicated to peripheral nerves. These nerves then carry taste information to the brain, where tastes are perceived.

The gustatory (i.e., taste) system is composed of three principal anatomical divisions: (1) taste buds within the gustatory epithelium of the tongue and the soft palate, which contain the taste sensory cells; (2) cranial nerves that connect these taste buds to gustatory areas of the brain stem; and (3) higher gustatory areas, such as the ventroposteriomedial nucleus of the thalamus and the gustatory cortex. The olfactory regions of the brain and brain stem are critical for processing taste information received from the periphery, helping to make sense of the signals the mouth detects. The taste buds, in contrast, are primarily concerned with the detection of taste stimuli. When tastants enter the mouth, they gain access to taste receptors on taste sensory cells within the taste buds, thus initiating the process of gustatory transduction.

Thousands of food compounds and other molecules have a taste. But the perceptual qualities elicited by those stimuli appear limited to just five: sweet, sour, bitter, salty, and umami (the savory taste of glutamate found in foods like meats and cheeses). For example, table sugar tastes sweet, but so do various no-calorie sweeteners, some amino acids and proteins, and even some heavy metals, such as lead. Why do such varied chemicals all taste sweet? The answer lies in the way taste stimuli are recognized. The gustatory system uses a few distinct populations of sensory cells in the oral cavity to detect tastants. Each sensory cell population is dedicated to the encoding of a different taste quality (i.e., sweet, bitter, etc.), and each of these sensory cell population expresses a different taste receptor or family of receptors.

The association of specific taste receptor types with distinct sensory cell populations dedicated to the coding of particular taste qualities dictates that all stimuli that can activate a particular taste receptor will elicit the same taste quality. In other words, activation of a "sweet" taste cell leads to the perception of sweetness, and activation of a "bitter" cell elicits bitterness. Which compounds can evoke these perceptions is dictated by the taste receptor(s) contained in different sensory cell types: The selectivity of these taste receptor(s) dictates that only certain tastants can elicit a particular taste quality. For example, the single taste receptor type expressed in "sweet" taste cells responds to numerous sugars and other natural and artificial sweeteners. Because all of these chemically diverse compounds can activate the taste receptor specifically expressed on "sweet" taste cells, all are perceived as sweet. Similarly, "bitter" taste cells employ a small number of taste receptors that together can recognize many hundreds of bitter-tasting compounds. Indeed, if your sweet taste receptor were expressed in "bitter" taste cells, table sugar would taste bitter to you. This entry covers the salt, sour, sweet, umami, and bitter tastes as well as the complexity of taste transduction.

Salt and Sour Taste

It is likely that there are two different mechanisms for transducing salt taste: one mechanism that is sensitive to attenuation by the drug amiloride, and one that is amiloride-insensitive. The partial blockage of salt taste by amiloride in many mammals suggests that there is more than one type of salt taste receptor and/or transduction mechanism (amiloride likely acts by blocking an ion channel, preventing the flow of sodium ions into the taste cell). However, although the amiloride-sensitive component of salt taste is prominent in rodents and some other mammals, it makes little contribution to salt taste in humans. Thus, humans may have only a single type of salt taste receptor. The receptor(s) for salt taste have not been identified, though several candidates have been proposed.

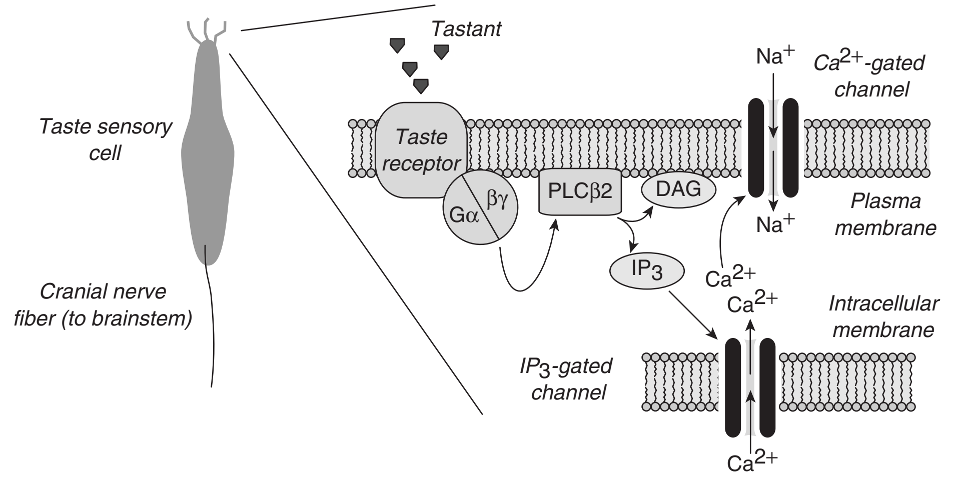


Figure 1 A General Model for the Transduction of Sweet, Bitter, and Umami Taste Stimuli; the Taste Receptor is Either a T1R (sweet, umami) or T2R (bitter)

The characteristics of the sour taste receptor also remain unclear. The identification of this receptor is hampered by the controversy as to what constitutes the primary sour stimulus. Some evidence supports free protons, which could either activate a sour taste receptor directly or could flux through an ion channel to depolarize the sour taste cell through a transfer of positive charge. Alternatively, uncharged acetic acid moieties may promote intracellular acidification and subsequent cellular activation. So, even though acetic acid and hydrochloric acid are both sour tasting, it remains unclear how such stimuli activate taste sensory cells.

In the case of both salt and sour taste, the relevant receptors are most likely ion channels (proteins that control the flux of ions across cell membranes) that are either gated by salts or acids or that allow the ionic stimulus to enter the cell. If this proves to be the case, then stimulus detection will be transduced into cellular depolarization through a single protein complex. In this fundamental way, the salt and sour receptors differ from the sweet, umami, and bitter receptors discussed next.

Sweet, Umami, and Bitter Taste

The sensory cells that mediate sweet, umami, and bitter tastes use a strikingly similar collection of molecules to recognize and transduce taste stimuli: (1) a receptor protein that sits within the plasma membrane of the sensory cell such that it can interact with both tastants and the transduction cascade within the cell; (2) an enzyme that produces a diffusible molecule (called a second messenger) upon receptor activation; and (3) an ion channel that opens in response to second messenger–mediated signals (e.g., Figure 1). The culmination of these cascades is the net movement of positive charge into the cell, depolarizing the cell membrane and promoting the release of neurotransmitter onto peripheral nerves.

Stimuli that elicit sweet, umami, or bitter taste perceptions are recognized by distinct families of receptors: The T1R receptors recognize sweet or umami tastants, and the T2R receptors recognize bitter tastants. Both T1Rs and T2Rs are members of the G protein coupled receptor (GPCR) superfamily, which includes most odorant, neurotransmitter, and hormonal receptors. As such, they share a number of structural and functional attributes with other GPCRs, including seven plasma membrane–spanning domains and the ability to couple to the transducing proteins called G proteins (see the following text). However, T1Rs and T2Rs differ significantly from each other and are only very distantly related.

The sweet taste receptor is composed of two separate proteins from the T1R family: T1R2 and T1R3. These two proteins are both required for normal responses to sweeteners, but act together as a single receptor complex. Binding of sweeteners to the T1R2:T1R3 receptor results in a change in the shape of the receptor, called a *conformational change*. This tastant-induced conformational change is transferred through the cell membrane to the cytoplasmic surface, where it stimulates a G protein complex composed of α, β, and γ subunits. One α subunit involved in sweet (and bitter and umami) taste transduction is known as α-gustducin, though others are also implicated. The principal β and γ subunits implicated in mammalian taste transduction are β_3 and γ_{13}. T1R2:T1R3 activation promotes the separation of the G protein α-subunit from the β and γ subunits, the latter of which exist in cells as a tightly bound complex. The $G\beta_3\gamma_{13}$ complex then interacts with the enzyme phospholipase C β2 (PLCβ2). PLCβ2 acts as an effector of taste receptor activation: upon interaction with $G\beta_3\gamma_{13}$, PLCβ2 generates two small signaling molecules, diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP_3). The role of DAG in taste cells is unclear, but the second messenger IP_3 promotes the release of calcium ions from cellular organelles into the cytoplasm of the taste cell. This rise in intracellular Ca^{2+} concentration is thought to open an ion channel in the cell membrane, promoting the influx of positively charged sodium ions, the depolarization of the cell membrane, and the release of neurotransmitter onto nerve fibers.

The transduction of umami tastants (e.g., the amino acids L-glutamate and L-aspartate) appears nearly identical to the transduction of sweeteners. Umami transduction relies on a similar GPCR-type taste receptor, which is composed of the T1R-family proteins T1R1 and T1R3 (the latter is shared with the sweet receptor). Umami transduction also employs G proteins, PLCβ2, the second messenger IP_3 and a Ca_{2+}-sensitive ion channel to transduce receptor activation into depolarization of the sensory cells. However, umami taste transduction differs from sweet taste transduction in a critical way. Umami-tasting stimuli activate a distinct cell population that specifically expresses the T1R1:T1R3 taste receptor, which is selective for certain L-amino acids. This subtle difference between sweet-sensitive and umami-sensitive cells—the taste receptors they express—illustrates the importance of the taste receptor in dictating which kinds of molecules (e.g., sugars vs. L-amino acids) can elicit a particular taste quality.

This principal holds true for bitter taste as well. Taste receptors in bitter-sensitive taste sensory cells are members of the T2R family. Mammals have 20 to 30 T2R genes. Most T2Rs are broadly tuned to many distinct compounds, though each T2R is likely the "best" receptor for a small number of compounds. Individual bitter-sensitive taste sensory cells appear to express most, if not all, T2Rs, indicating that these cells function as broad sensors for hundreds or thousands of compounds. The coexpression of multiple T2Rs in the same sensory cell also has the consequence that all compounds that activate T2Rs taste alike. In other words, with the exception of intensity, humans and other mammals cannot differentiate one bitter-tasting substance from another. However, it is not uncommon for individuals to possess one or more nonfunctional T2R genes, leaving those people ageusic, or "taste-blind" to certain taste compounds. T2Rs utilize the same basic molecular mechanisms as do T1Rs to transduce taste stimuli: A stimulus induced conformational change in the T2R receptor elicits a cellular depolarization that is dependent on G proteins, PLCβ2, IP_3, and an ion channel. Thus, with the principal exceptions of receptor and sensory cell population, sweet, umami, and bitter taste transduction is remarkably similar.

The Complexity of Taste Transduction

The detection and transduction of taste stimuli is complex. The use of separate sensory cell populations to encode different taste qualities provides an opportunity to distinguish taste stimuli that carry very different information about the nutrient content and safety of food. The several distinct taste receptors and their mutually exclusive expression across sensory cell types afford a degree of selectivity that pairs stimulus type with perceptual quality. The varied transduction mechanisms brought to bear in the gustatory system reflect this stimulus and receptor diversity, allowing receptor stimulation to be efficiently coupled with cellular activation.

Interestingly, the impact of these receptors and transduction molecules are not restricted to the gustatory system: Some cells in the gastrointestinal tract use the same proteins to sense nutrients and other compounds. Thus, elucidating the complex processes involved in the recognition of taste stimuli will impact both our ingestive choices and our metabolic health.

Steven D. Munger

See also Olfactory Receptors and Transduction; Taste; Taste: Genetics of; Taste Adaptation; Taste and Food Preferences; Taste Stimuli: Chemical and Food; Taste System Structure; Taste Thresholds and Intensity

Further Readings

Bachmanov, A. A., & Beauchamp, G. K. (2007). Taste receptor genes. *Annual Review of Nutrition, 27,* 389–414.

Boughter, J. D., Jr., & Bachmanov, A. A. (2007). Behavioral genetics and taste. *BMC Neuroscience, 8*(Suppl. 3), S3.

Breslin, P. A., & Spector, A. C. (2008). Mammalian taste perception. *Current Biology, 18*(4), R148–R155.

Chandrashekar, J., Hoon, M. A., Ryba, N. J., & Zuker, C. S. (2006). The receptors and cells for mammalian taste. *Nature, 444,* 288–294.

Rozengurt, E., & Sternini, C. (2007). Taste receptor signaling in the mammalian gut. *Current Opinion in Pharmacology, 7,* 557–562.

Scott, K. (2005). Taste recognition: Food for thought. *Neuron, 48,* 455–464.

TASTE STIMULI: CHEMICAL AND FOOD

In the lexicon of the consumer, taste refers to the entire sensory experience of a food or beverage. Thus, it includes visual sensation. Is the cola beverage brown like it should be? Are the grapes the right size and without deformities? It also includes auditory sensation. Does the apple crunch when biting into it? Does the carbonated beverage fizz when pouring? And it includes tactile sensation. Is the banana firm or is it soft? Is the milk shake thick and rich, or is it thin and watery? Olfactory sensation is a major contributor to food and beverage perception and is more confused with taste than any of the foregoing. But really, taste is a gustatory sensation. It emanates from specialized sensory cells of the taste buds located on the tongue and some other parts of the oral cavity. All of the other four senses contribute significantly to the overall perception of foods and beverages but are not taste. Only gustatory sensation is taste.

Gustation is constituted of the four well-known primary tastes: sweet, bitter, sour, and salty, and one lesser-known taste known as savory, but often referred to as *umami,* the Japanese word for delicious. The stimuli initiating these tastes are all chemical compounds occurring naturally in foods and beverages. For the case of sweet taste, due to health issues associated with diabetes and obesity, noncaloric sweeteners are now in use, as well as the common caloric sweeteners.

Many protocols have been employed in the sensory testing of chemical compounds. In order for a compound to have taste, it must be soluble in water and saliva. If a solid chemical is tasted, its taste will be dependent on its rate of dissolution in saliva, a variable that is difficult to control (i.e., dependent on solid particle size, crystal form, level of agitation of the solid/saliva mixture, etc.), and therefore sensory testing is not generally carried out with solid samples. At the same time, however, in real life, consumers often experience solid stimuli in the course of normal eating. Nonetheless, in almost all cases, sensory studies on gustatory stimuli are carried out on prepared solutions with known concentrations. In other cases, where this is not possible, stimuli concentrations are determined by chromatographic or spectroscopic methods. This entry describes sweet-tasting, bitter-tasting, umami-tasting, sour-tasting, and salty-tasting stimuli, as well as stimuli exhibiting mixtures of primary taste qualities.

Sweet-Tasting Stimuli

Sweet taste is thought to be a sense that evolved for identification of nutritive foods and beverages. Fruits and vegetables contain nutritive sugars and polyols of which sucrose, glucose, and fructose are the most common. In addition, a great many other low molecular weight carbohydrates are known and all exhibit a sweet taste, although some exhibit bitterness as well. Because sucrose is naturally present in many fruits and vegetables, it is produced as a pure substance from sugar cane and sugar beets, on a very large scale, with world production at approximately 150 million metric tons per annum. Because glucose and fructose are also present naturally in fruits and vegetables, they are produced from starch on a large commercial scale, both as pure carbohydrate sweeteners and as blends commonly referred to as high fructose starch syrup (HFSS), or high fructose corn syrup (HFCS) when prepared from corn starch. The most common starch based sweeteners are HFSS-42 (42% fructose) and HFSS-55 (55% fructose).

Over the last century, many noncaloric sweeteners have been commercialized, principally synthetic but also natural. In addition, several thousand synthetic and natural noncaloric sweeteners have been identified in the laboratory. These sweeteners vary significantly in quality of taste (i.e., presence of “off” tastes such as bitter or cooling, rates of sweetness onset/dissipation, and levels of sweetness desensitization) as well as in sweetness potency (P), usually expressed as a multiple of sucrose sweetness. Early in the 20th century, the primary interest in noncaloric sweeteners was to address the special needs of diabetics. Thus, saccharin (P = 180) was commercialized in the United States as the first product of Monsanto Chemical Company. Later, interest in noncaloric sweeteners intensified due to the increasing prevalence of obesity. Thus, synthetic sweeteners commercialized in the second half of the 20th century included salts of cyclamic acid (P = 23 / sodium salt), commonly known as cyclamates, aspartame (P = 140), acesulfame (P = 76) as a potassium salt, sucralose (P = 500), and neotame (P = 10,000). Two natural noncaloric sweeteners were also commercialized, both with generally regarded as safe (GRAS) status in the United States. The first of these is erythritol (P = 0.65) and the second, rebaudioside A (P = 250), is a secondary plant metabolite from the Paraguayan plant *Stevia rebaudiana* (Bertoni).

Although most sweet stimuli are nutritive sugars or safe synthetic or natural sweeteners, not all sweet substances are safe for consumption. Many salts of lead are sweet and were in general usage in paint formulations in the first half of the 20th century. Paint chips in old housing often contained lead salts and, in some cases, led to toxic effects in children. Lead-based paints are now banned in developed countries. Another sweet-tasting and toxic chemical compound is ethylene glycol, which is commonly used as an antifreeze in automobile cooling systems. Careless use has resulted in accidental poisoning of dogs, and so it must be used with care.

Bitter-Tasting Stimuli

Bitter taste is thought to be a sense that evolved to enable avoidance of toxic substances in the environment because a high percentage of toxic chemicals are bitter in taste. Some bitterants are not toxic, however, and are naturally present in foods (e.g., caffeine in coffee, tea, and chocolate, naringin in grapefruit) and some bitterants (e.g., quinine in tonic water) are intentionally added to foods and beverages, thus providing a key component of the flavor profile. Others (e.g., sucrose octaacetate) are used primarily for laboratory studies to explore species specificity and mechanism of action. In addition, at least one bitterant (i.e., denatonium chloride) is intentionally added to some toxic consumer products to render them so unpalatable that children will not ingest them.

Umami-Tasting Stimuli

Umami taste is thought to have evolved to enable identification of protein-rich foods and beverages. Proteins are constituted of amino acid building blocks and many of the 20 amino acids in proteins are essential for life. The principal umami stimulants are the salts of glutamic acid (e.g., monosodium glutamate = MSG) and aspartic acid, and MSG is a common food additive that provides umami taste to foods. Some nucleotide monophosphate salts (e.g., monosodium inosine monophosphate = IMP) are said to be synergistic with MSG; however, they actually have no umami taste, but rather enhance the umami taste of glutamate and aspartate salts.

Sour-Tasting Stimuli

All acids are sour in taste. Sour taste, like bitter taste, likely evolved to enable avoidance of harmful substances. However, many foods and beverages (e.g., fruits and fruit juices) exhibit sour taste in combination with sweet taste to provide good sources of nutrition as well as great quality of taste. Phosphoric acid, a mineral acid, is the acid present in some beverages (e.g., cola) and in its neutral form is naturally present at significant levels in all living organisms. As an example, complex phosphate salts are the major component of bone. Fruits and vegetables contain many organic acids (e.g., citric acid, malic acid, tartaric acid, fumaric acid, and acetic acid), and these are often used in finished food and beverage formulations. For fresh fruits and vegetables, sourness is a metric of ripeness. Before ripening, sourness is high and the overall taste is unpleasant. However, on ripening, sourness decreases, sweetness increases, and the overall sensory experience is pleasant. Sour taste can be due to any acid. However, as far as sourness goes, all acids are not created equal. At the same pH, organic acids are notably more sour than mineral acids.

Salty-Tasting Stimuli

Salty taste is the taste of sodium chloride, lithium chloride, potassium chloride, and some other salts. Not all salts are salty in taste. Many show a mixture of tastes, with bitter taste being commonly present and where salty taste is only a minor component. Salty taste is thought to be a sense that evolved to enable mammals to find mineral nutrients including sodium and chloride as well as potassium, all of which are essential for life. All living cells are dependent on sodium, potassium, calcium, chloride, and a few other ionic species and thus ingestion of these minerals is a critical component of any healthy diet.

Stimuli Exhibiting Mixtures of Primary Taste Qualities

Many gustatory stimuli do not exhibit taste of a single gustatory modality (i.e., sweet, bitter, sour, etc.). For example, it is well known that although saccharin is perceived to be sweet, it is also perceived as bitter by many people. Likewise, potassium chloride, the common salt substitute used by many salt-sensitive hypertensive people, exhibits both salty and bitter tastes. It is now understood that such mixed sensation stimuli initiate these effects by activation of multiple types of taste bud cells (e.g., sweet- and bitter-sensitive taste bud cells for saccharin).

Grant E. DuBois

See also Taste; Taste and Food Preferences; Taste Thresholds and Intensity

Further Readings

Ackerman, D. (1990). *A natural history of the senses.* New York: Random House.

Beets, M. G. J. (1978). *Structure-activity relationships in human chemoreception.* London: Applied Science.

Coren, S., & Ward, L. M. (1989). *Sensation and perception* (3rd ed.). New York: Harcourt Brace Jovanovich.

DuBois, G. E. (2000). Non-nutritive sweeteners. *Encyclopedia of food science & technology* (pp. 2245–2265). New York: John Wiley & Sons.

DuBois, G. E., DeSimone, J., & Lyall, V. (2008). Chemistry of gustatory stimuli. In D. Smith & S. Firestein (Eds.), *The senses* (Vol. 2, pp. 28–74). London: Elsevier.

Moskowitz, H. R. (1983). *Product testing and sensory evaluation of foods: Marketing and R&D approaches.* Westport, CT: Food & Nutrition Press.

Shallenberger, R. S. (1993). *Taste chemistry.* London: Blackie Academic & Professional.

TASTE SYSTEM STRUCTURE

Evolution equipped living organisms with the ability to detect and respond to chemicals in their environments for the purposes of communication, reproduction, and nutrient detection. Many airborne chemicals are detected by the sense of smell, whereas chemicals in solution that are subject to ingestion are sensed and evaluated by the gustatory (taste) system. The sense of taste acts as a final gating mechanism to the internal milieu: The taste of a substance plays a large role in determining whether or not it will be ingested. Thus, taste critically guides the selection and intake of foods, intimately influencing our health and well-being. This entry describes the structure of the taste system, emphasizing the roles of the various structures in taste processing.

Humans can categorize taste sensations into one of five different general perceptual categories: sweet, salty, sour, bitter, and umami. *Umami* is a Japanese word that refers to the "savory" or "meaty" taste of certain amino acid stimuli, such as monosodium glutamate (MSG). These different perceptions begin when taste chemicals released from food in the mouth bind with specialized receptors for these chemicals that are linked to the neural pathway underlying taste. Such receptors are known simply as taste receptors. It is noteworthy that when humans describe and talk about the taste of a food they are usually referring to its flavor, a concept that involves smell, tactile (touch and texture) sensations, temperature, and also taste. Yet in precise terms, "taste" originates from the interaction of chemicals dissolved in saliva with taste receptors.

From what researchers currently know, taste receptors are composed of proteins that lie on the surface of "taste cells." In general, a taste cell is a specialized epithelial (skin) cell in the mouth that harbors taste receptors. These cells become "activated" when a taste chemical binds to a receptor they might express. More specifically, the binding between a taste chemical and its receptor results in a cascade of molecular events inside a taste cell, culminating in a change of the electrical charge across the cell's membrane. This change in electrical potential provides the basis of a signal sent to the brain indicating that a taste stimulus is present and what kind of stimulus it is. In humans and other mammals, taste cells are clumped together in onion-shaped structures called taste buds, each housing 50 to 100 or so cells, including taste cells and supporting cells. Taste buds are found in various regions of the oral cavity, including on the tongue and the soft palate on the roof of the mouth. Taste buds are housed in papilla, which on the tongue are the small protrusions (or bumps) that give this organ its velvety appearance. In most vertebrates, there is continual turnover and replacement of taste cells throughout life. Taste cells have been shown to have a lifespan of 9 to 10 days on average.

It is noteworthy that taste receptors sensitive to different types of taste chemicals (i.e., sweet, salty, bitter) are found in the same taste-cell-containing regions of oral epithelia. Receptors for a single class of taste stimulus are not restricted to specific locales, as has been depicted by the classic "taste map" of the tongue. Although sometimes still finding its way into modern textbook chapters on taste, the tongue map, which is based on old data, has long been considered by taste researchers to be incorrect. It is now known that taste receptors for pleasurable sweet-tasting stimuli are found in taste bud cells located on the front and rear of the tongue, although the classic tongue map would show that sweet can only be detected on the tongue's tip. The outdated tongue map also depicts bitter taste being mediated exclusively by taste receptors on the rear of the tongue. However, the application of bitter stimuli to the front of the tongue produces strong activation of taste circuitry in the brain, indicating that there are indeed bitter taste receptors on the front of the tongue. What is more, it has shown in rodent experiments that disrupting the cranial nerve carrying taste input from the front of the tongue to the brain results in an impairment of a rat's ability to recognize the taste of bitter stimuli in certain behavioral tasks, suggesting that the front of the tongue is critically involved in bitter taste perception.

Taste receptors and taste cells are involved in only the earliest stages of taste processing by the nervous system. Fibers of cranial nerves carry information from taste receptors into a vast network of neurons and nuclei downstream in the central nervous system, and it is here in the brain that messages from taste receptors are read out and converted to perceptual and behavioral responses. Taste information arrives at the central nervous system at the level of the brain stem, at a structure known as the *nucleus of the solitary tract* (abbreviated NTS for its Latin name: nucleus tractus solitarius). Nerve fibers from the different taste-cell-containing regions of the mouth terminate in NTS in a loosely "mapped" fashion. The most forward region of gustatory NTS receives input from nerve fibers projecting mainly from taste cells on the front of the tongue, whereas the most rearward portion of gustatory NTS receives a majority of its input from nerve fibers innervating taste cells on the back of the tongue. In other words, there is a systematic topographic arrangement to the pattern of taste nerve projections and their terminations in the brain. Topography appears to be a common organizing principle in sensory systems, as also seen, for example, in the somatosensory system, where neighboring regions of the body are systematically represented along neighboring regions of the somatosensory cortex.

There have been many studies of the taste response properties of individual taste-sensitive neurons in NTS (and in other brain taste nuclei as well). These studies have shown that individual taste neurons can display different "tuning orientations" toward taste stimuli. For example, some sweet-sensitive NTS taste neurons respond rather selectively to only sweet stimuli, being only weakly or not at all activated by particular nonsweet stimuli. In contrast, other NTS neurons have been found that show very strong responses to, for example, the sweet taste of the sugar sucrose and also to the salty taste of sodium chloride (table salt). Basically, there appears to be a wide mix of different tuning orientations among NTS neurons. It is this mix of tuning profiles among taste neurons that has posed a challenge to gustatory neurobiologists attempting to unravel the mechanism by which these cells signal information about different taste stimuli. If taste neurons exist that are responsive to multiple kinds of taste stimuli, such as sweet and salty, how does the taste system correctly "read out" the identity of any single taste stimulus without confusion? We know the brain accomplishes this task with ease: You can intuit that you always recognize and easily distinguish the tastes of sugar, salt, or any other taste stimulus for that matter. Although some researchers have proposed that perceptual recognition of unique taste stimuli is achieved by the brain attending to the overall activity of a group of taste neurons, a clear solution to this problem remains to be fully worked out.

In rodents, taste information from NTS is projected to the parabrachial nucleus (PbN) of the pons located in the midbrain, which routes taste input to higher-order brain areas involved in perceptual integration, feeding, motivation, and reward. It is noteworthy that in primates and humans the PbN appears to have been bypassed as a waypoint for taste processing, with taste input reaching higher-order nuclei via direct projections from NTS. Some of the higher-order regions of the brain involved in taste processing include the gustatory cortex (the insula), the hypothalamus (associated with the control of feeding), and the amygdala (associated with affective/emotional processing). Interestingly, the cortical and forebrain structures involved in taste maintain descending connections with taste circuits in PbN and NTS. Electrical stimulation or anesthetization of higher-order gustatory areas in the brain has been shown to change the way that the NTS and PbN respond to taste input. Descending control exerted from structures involved in affect (amygdala) or hunger (hypothalamus) on nuclei that reside early in the ascending taste pathway could possibly constitute a mechanism to allow the gustatory system to dynamically adjust its overall sensitivity in the face of the organism's current nutritional needs, behavioral status, or experience.

Christian H. Lemon

See also Taste; Taste Adaptation; Taste Receptors and Transduction; Taste Thresholds and Intensity

Further Readings

Jones, L. M., Fontanini, A., & Katz, D. B. (2006). Gustatory processing: A dynamic systems approach. *Current Opinion in Neurobiology, 16*(4), 420–428.

Lemon, C. H., & Katz, D. B. (2007). The neural processing of taste. *BMC Neuroscience, 8*(Suppl. 3), S5.

Smith, D. V., & Margolskee, R. F. (2001). Making sense of taste. *Scientific American, 284*(3), 32–39.

Spector, A. C., & Travers, S. P. (2005). The representation of taste quality in the mammalian nervous system. *Behavioral and Cognitive Neuroscience Reviews, 4*(3), 43–191.

TASTE THRESHOLDS AND INTENSITY

Perceptual scientists often study sensory responses at two levels: responses to weak and to strong stimuli. The simplest form of weak stimulus study is the measurement of absolute detection threshold. In taste research, this is the lowest level of a sapid chemical that can be distinguished from pure water. Threshold concentrations for taste stimuli range from micromolar (μmol) levels for many toxic compounds, such as quinine (a commonly experienced bitter drug and beverage flavorant) to low millimolar (mmol) ranges for NaCl (a salty stimulus), citric acid (a sour stimulus), and glutamate (a savory stimulus), and even to higher millimolar range for some sugars (sweet stimuli). The simplest form of a strong stimulus study is the measurement of sensation magnitude with questions like, "How strong is the salty taste?" This entry discusses the relation between taste thresholds and perceived intensity of taste stimuli.

Usually these two types of measures correlate because they are encoded by the same peripheral physiological systems. But what do lay people appreciate of the differences between these two sensory responses? Someone with a high fever or a migraine headache might declare that their sensitivity to stimuli, such as lights, sound, touches, and odors, is high. By this they could mean that they are either bothered by barely detectable stimuli that would otherwise not be noticed or that they are pained by intermediately intense stimuli that would otherwise seem only weak to moderate in strength, or most likely both. The general habit of linking together our sensitivities to threshold-level stimuli and to suprathreshold stimuli is born of a common understanding that absolute threshold sensitivity and the magnitude of suprathreshold responses are related. Our life experiences confirm the relationship between these two.

But these two measures are often not correlated for taste stimuli. For example, someone who can detect extremely low levels of salt in water, lower than most others can, may find very high concentrations of salt in water to be only weakly salty. Or conversely, someone who requires relatively high concentrations of quinine to notice it might report that moderately stronger concentrations of quinine are overwhelmingly bitter. Observations such as this, which indicate a lack of correlation between threshold and suprathreshold measures, suggest that these measures may involve different cellular and molecular processes.

The popular understanding of the relationship between our perceptual responses to weak and strong stimuli contains two assumptions. First, we implicitly assume that detection and suprathreshold ratings of stimuli are based upon the same systems, a correct assumption for organs such as the eye, the ear, and the skin. Second, we assume that the sensory systems that detect a weak stimulus grow in their response with stimulus strength similarly for different people. However, at the absolute threshold for taste stimuli, detection level sensations are not usually accompanied by a quality of taste. Rather, it is merely the detection that something rather than nothing is present in the water. At this low level of stimulation in the human mouth, we cannot be certain of the identity of the receptor system that detects the stimulus. It is possible that detection at these low levels is not based upon activation of the taste system, but rather a somatosensory or osmotic input provides the signal.

Taste transduction involves many different receptors as well as the receptors' down-stream signaling components within the receptor cell, such as G proteins, second messengers, and gated ion channels, all of which exist in multiple forms (they are polymorphic within the population) that possess different functional properties. Therefore, the absolute sensitivity of a receptor cell and the rate at which cellular activation grows with stimulus strength may vary from person to person as a function of the different variants of the transduction proteins a person possesses. A particular insensitive receptor variant in the taste cell may make a person relatively insensitive to the stimulus because detection begins with the receptor response. But the forms of downstream signaling proteins may also give the cell a very high amplification of signal that results in the rapid growth of cellular activation with increases in stimulus strength. This would result in a person with high detection thresholds (who was insensitive to low concentrations), but who perceived very strong taste from slightly higher suprathreshold concentrations. The complicated genetics of taste transduction can account for the independence of these different perceptual variables.

There are also examples of taste stimuli for which people's absolute sensitivity predicts their perceived taste intensity with higher stimulus concentrations. The single best example of this is the bitter taste of the antithyroid drug phenylthiocarbamide (PTC). Absolute sensitivity to PTC is determined by the variant forms of the PTC taste receptor (TAS2R38) that people possess. There is a common form of the receptor that is very sensitive to PTC, and a common form that is insensitive.

People who possess two copies (homozygous) of the sensitive form of the receptor detect PTC at micromolar concentrations. They also perceive higher suprathreshold concentrations of PTC as intensely bitter, sometimes overpoweringly so. Those who are homozygous for the two insensitive forms of the receptor can only detect PTC at concentrations 1,000 times stronger, and high suprathreshold concentrations are usually only perceived as weakly bitter. This correlation between detection threshold sensitivity and perceived intensity of suprathreshold concentrations of PTC is explained by the impact of the receptor variants on perception of PTC. Almost all of the variation in PTC taste responses is determined by the two common forms of the receptor. Within the subset of people who are strong responders to PTC, however, there is variation in their bitter response that is most likely determined by variations in other transducing components that affect amplification processes.

The absolute sensitivity of the taste system is important for detecting the presence of nutrients and toxins in the foods we eat. The magnitude of taste at high concentrations of stimuli is important for assessing the levels of these nutrients and toxins in foods. The mechanistic explanation for their perceptual independence within a population remains to be elucidated, but will likely involves the genetics of receptor systems as detectors and their signal amplification systems within the receptor cells.

Paul A. S. Breslin

See also Taste; Taste: Genetics of; Taste Receptors and Transduction

Further Readings

Bartoshuk, L. M. (1978). The psychophysics of taste. *American Journal of Clinical Nutrition, 31,* 1068–1077.

Bartoshuk, L. M. (2000). Comparing sensory experiences across individuals: Recent psychophysical advances illuminate genetic variation in taste perception. *Chemical Senses, 25,* 447–460.

Breslin, P. A., & Spector, A. C. (2008). Mammalian taste perception. *Current Biology, 18,* R148–R155.

Keast, R. S., & Roper, J. (2007). A complex relationship among chemical concentration, detection threshold, and suprathreshold intensity of bitter compounds. *Chemical Senses, 32,* 245–253.

TEMPERATURE PERCEPTION

In contrast to the wealth of perceptual experiences associated with visual and auditory stimulation, the range of sensations elicited by changing the temperature of the skin is very limited. In response to thermal stimulation, people generally report that their skin has been warmed or cooled and this is quantified in terms of intensity and duration. The hedonic nature (i.e., pleasantness/unpleasantness) of the thermal stimulus may also be attended to (e.g., the pleasurable sensations associated with standing by a warm fire on a cold day and at extreme temperatures sensations of pain are evoked).

When objects are held in the hand, changes in skin temperature can be used to assist in identification and discrimination. These cues become particularly important when objects must be identified in the absence of vision, such as when reaching for objects in the dark or when visual attention is directed elsewhere, for example, when driving. The thermal cues that are used to assist in identifying an object arise from changes in skin temperature that occur when the object is grasped. The thermal properties of the object and the initial temperatures of the skin and object all determine the rate at which heat is conducted out of the skin or object on contact. As the resting temperature of the skin is typically higher than the temperature of objects encountered in the environment, it is the decrease in skin temperature on contact that is generally used to help identify an object. This entry describes thermal receptors and thresholds, spatial and temporal aspects of perception, and the contribution of thermal cues to object identification and discrimination.

Thermal Receptors

The ability to perceive innocuous temperatures depends primarily on two kinds of receptors found in the skin known as cold and warm thermoreceptors. Cold receptors respond to decreases in temperature over a range of 5 to 43° Celsius and discharge most vigorously at skin temperatures around 25°C. In contrast, warm receptors signal that skin temperature has increased and are most responsive at approximately 45°C. When the temperature of the skin is maintained at 30 to 36°C, which is the normal range for skin temperature, both types of receptor are spontaneously active, but there is no awareness of cold or warmth.

The density of thermoreceptors in the skin has been analyzed in terms of warm and cold spots, which are identified perceptually by placing small warm and cold stimulators on the skin and recording when a stimulus is detected. It is assumed that thermal spots mark the receptive fields of underlying thermoreceptors, which are probably free nerve endings. Barry Green and his colleagues have analyzed maps of these spots, which are only a few millimeters in diameter, and found that warm and cold spots are independently distributed, that cold spots outnumber warm spots, and that the density of spots varies at different sites on the body. For example, on the forearm there are patches of skin several square centimeters in area (i.e., the size of a postage stamp) within which warm stimuli are not detected at all.

A further difference between cold and warm thermoreceptors is the conduction velocities of the afferent fibers that arise from the receptors. Cold afferent fibers are much faster than warm afferent fibers, with conduction velocities of 10 to 20 meters per second (m/s) as compared to 1 to 2 m/s for warm fibers. As would be expected from these differences in conduction velocities, the reaction time for the development of cold sensations is significantly shorter than that for warmth.

The skin also houses thermally sensitive pain receptors known as thermal nociceptors that respond to noxious or harmful temperatures and signal to the central nervous system that tissue damage is imminent and that the affected body part should be withdrawn from the thermal source (e.g., a finger on a hot plate). These receptors are active when the temperature of the skin falls below 15 to 18°C or rises above 45°C, and when they are activated there is a fairly sharp change in the character of thermal sensation to one of pain. Although the thresholds for heat- and cold-sensitive nociceptors are usually described as being greater than 45°C and less than 15°C, in some individuals mild cooling (25–31°C) and warming (34–40°C) of the skin can evoke sensations of burning and stinging as well as innocuous sensations of cold and warmth. In addition to thermoreceptors and nociceptors, there are mechanoreceptors in the skin whose activity is affected by changes in skin temperature. It is generally accepted that these mechanoreceptors that respond to deformation of the skin do not affect thermal sensation.

Advances in understanding the cellular and molecular mechanisms involved in sensing warmth and cold are contributing to a better comprehension of the processes involved in converting thermal information into chemical and electrical signals in the peripheral nervous system. The process of thermal sensation begins with specific receptor proteins located within the free nerve endings in the skin. A class of sensory ion channels known as temperature-activated or thermo transient receptor potential (TRP) ion channels has been identified in sensory nerve endings that respond to thermal stimuli. These temperature-activated TRP channels have been cloned and characterized, and individual thermo TRPs have been shown to be specialized to detect distinct temperature ranges.

The neural pathways subserving temperature perception have been difficult to identify conclusively, but considerable progress has been made over the past decade. Thermoreceptors innervating the skin terminate in the superficial spinal dorsal horn where a population of spinothalamic neurons is specific to thermal inputs and morphologically distinct. These neurons have thresholds close to normal skin temperature and have linearly graded responses to innocuous cooling and an ongoing discharge that is inhibited when the skin is warmed. Their activity is correlated with the ability to discriminate between two temperatures. Neurons that are selectively excited by warm stimuli are relatively rare. Spinothalamic neurons in the dorsal horn project to the thalamus via the crossed lateral spinothalamic tract, and in the human brain the neurons terminate in the ventral caudal nucleus of the thalamus, which contains both thermoreceptive- and nociceptive-specific neurons. Imaging (positron emission tomography) data from humans indicate that neurons in the ventral thalamus project topographically to the dorsal margin of the middle/posterior insular cortex, an area associated with autonomic and limbic function.

Thermal Thresholds

The ability to perceive changes in temperature depends on several factors, including the amplitude and rate of temperature change, the baseline temperature of the skin, and the site on the body stimulated. The thermal sensory system is extremely sensitive to small changes in temperature, and on the glabrous (hairless) skin at the base of the thumb people can perceive a difference of 0.02 to 0.07°C in the amplitudes of two cooling pulses or 0.03 to 0.09°C in two warming pulses delivered to the skin. The absolute threshold for detecting a change in skin temperature is larger than the threshold for discriminating the difference between two cooling or warming pulses. When the skin at the base of the thumb is maintained at 33°C, the threshold for detecting an increase in temperature is 0.20°C and is 0.11°C for detecting a decrease in temperature.

If the temperature of the skin changes very slowly, for example at a rate of less than 0.5°C/minute, then a person can be unaware of a change of up to 4 to 5°C, provided that the temperature remains within the neutral thermal zone of 30 to 36°C. If temperature changes at a more rapid rate, such as at 0.1°C/second, then small increases and decreases in skin temperature can be detected. However, there is very little additional effect on either warm or cold thresholds if the rate of temperature change is faster than 0.1°C/s. The higher thresholds that occur with slower rates of temperature change can be accounted for by adaptation, that is, the loss of responsiveness of primary afferent units to thermal stimulation with time.

The baseline temperature of the skin also influences the ability to detect changes in temperature. At low skin temperatures of around 28°C, the threshold for detecting a warm stimulus is large (1°C), whereas the threshold for a cold stimulus is small (0.1°C). As the temperature of the skin increases, the warm threshold decreases and the cold threshold increases, and by 40°C the warm threshold is 0.2°C and the cold threshold is 1.1°C.

There is a 100-fold variation across the body in its sensitivity to changes in temperature, with the cheeks and the lips being exceptionally sensitive and the extremities, particularly the feet, being relatively insensitive. A common finding in many studies of thermal thresholds is that despite the variability in thresholds, all body regions are more sensitive to cold than to warm stimuli. In general, warm thresholds are twice as large as cold thresholds, and the better a site is at detecting cold the better it is at detecting warmth.

Spatial Aspects of Perception

Spatial Summation

The spatial features of a thermal stimulus, such as its area and shape, are not clearly perceived, and changes in intensity within a region of thermal stimulation are often not detected. However, if the skin is warmed or cooled as a result of touching an object, then the location and dimensions of the thermal stimulus can be determined on the basis of the tactile input. The inability to resolve variations in temperature within an area of stimulation occurs because the thermal sensory system summates thermal stimulation across the skin surface. This increases the detectability of small changes in temperature that occur over a large surface area, such as the torso or arms, which is important in the regulation of body temperature. Small increases or decreases in skin temperature integrated over a large area will result in a considerable change in thermal load for the thermoregulatory system, which controls core body temperature via vasodilation, and if necessary sweating when hot, and vasoconstriction and shivering when cold.

For the thermal senses, changing the spatial extent of stimulation primarily affects the perceived intensity of the stimulus, rather than its perceived size. This is an unusual feature of a sensory system, and would be equivalent in vision to a stimulus whose area increased being perceived as brighter rather than simply appearing bigger. For warm stimuli, the contribution of the size of the area stimulated to estimates of warmth diminishes as the intensity of the stimulus increases and approaches the heat pain threshold. When the skin is cooled, the perceived intensity of the cold stimulus grows at a nearly constant rate as a function of contact area, regardless of the degree of cooling.

An increase in the area of stimulation also results in the thermal stimulus becoming more detectable. When the area of skin that is warmed is doubled, the threshold is halved. This reciprocity between area and threshold does break down as the intensity of the warm stimulus approaches the threshold for pain. The same relation between intensity and the area of stimulation holds for the perception of cold, although area does not play as prominent a role in determining cold thresholds. For cooling stimuli, the threshold can be halved with a four-fold increase in the area of skin cooled.

Spatial summation not only occurs at a single site of increasing area, but has also been demonstrated when two sites symmetrically located on opposite sides of the body are stimulated simultaneously. There is summation of both threshold and suprathreshold warm and cold stimuli when they are presented bilaterally. This means that if the same thermal stimuli are presented simultaneously on the left and right forearms, the threshold calculated is significantly smaller than that measured when a single site on one arm is stimulated. However, there is no change in threshold if the sites are asymmetric (e.g., the forehead and the contralateral hand). These findings suggest that thermal spatial summation is centrally, rather than peripherally, mediated as the sites of stimulation over which it occurs are spatially distinct but functionally related. It has also been proposed that variations in thermal sensitivity across the skin could significantly influence spatial summation and that these variations could account for much of the increase in sensitivity measured as the area of skin stimulated increases.

Spatial Acuity

The thermal sensory system is poor at localizing the site of stimulation on the body and at spatially differentiating two thermal stimuli, as would be predicted for a sense that displays such rich spatial summation. When tactile cues are eliminated by using noncontact thermal stimuli, such as radiant heat, localizing the site of stimulation is very poor, particularly at low levels of stimulus intensity. On large surface areas, such as the torso, people occasionally mislocalize warm stimuli presented on the back as being displayed on the front of their torso. However, as the intensity of warm stimuli approaches the pain threshold, the ability to localize them improves considerably, as now there is a relative urgency in localizing the point of stimulation before tissue damage occurs.

Although simultaneous tactile cues can help localize thermal stimulation, the interaction between thermal and mechanical inputs can lead to mislocalization of thermal sensations when adjacent parts of the skin are differentially stimulated. B. G. Green described a thermal illusion involving the hand in which the thermal sensations arising from the middle finger changed as a function of the sensations experienced at the two adjacent fingers. When the index and ring fingers were placed on cold (or warm) thermal stimulators and the middle finger was placed on a thermally neutral stimulator, cold (or warmth) was felt on all three fingers. This referral of thermal sensations required equivalent tactile experiences on the fingers, in that it did not occur when the middle finger was held above the stimulator. These findings indicate that information about the locus of stimulation acquired through the tactile sense is made available to the thermal senses, even though these two sources of information are conveyed via distinct neural pathways.

Temporal Aspects of Perception

The temporal aspects of thermal stimuli can have a profound effect on perception. Similar to the spatial summation effects previously described, the duration and intensity of a thermal stimulus can trade so as to preserve constant a threshold or suprathreshold sensation. This property of temporal summation has mainly been investigated with warm stimuli presented on the face, where it has been observed that for durations up to 1 s the warm threshold varies in nearly inverse proportion to the duration of infrared irradiation. This means that the intensity needed to detect a warm stimulus is halved if the duration is doubled. On the face, a thermally sensitive site, 1 s was the critical duration; at shorter durations, intensity and duration traded but not proportionally; and at durations longer than 1 s there was no further influence of duration on the warm threshold. In the context of other sensory modalities, the critical duration of 1 s in the thermal senses is considerably longer than the 0.1 to 0.2 s reported for vision and hearing, and so reflects the rather sluggish character of the thermal senses.

The decrease in responsiveness to thermal stimulation as a result of continuous exposure to the stimulus is referred to as adaptation and is a frequently experienced phenomenon. For example, the sensation of warmth that is aroused when the hand is plunged into warm water gradually diminishes and may completely disappear within a few minutes. The skin adapts to both warm and cold stimuli over time, and for the forearm there is complete adaptation within about 25 minutes for temperatures between 28 and 37.5°C. The rate at which adaptation occurs is very rapid for temperatures close to that of the skin and much slower for more extreme temperatures.

Contribution of Thermal Cues to Object Identification and Discrimination

Changes in skin temperature contribute to the sensory information about an object in contact with the skin. The sensation of warmth or coldness that occurs when the hand makes contact with an object is a transient heat conduction phenomenon. As the temperature of the hand is generally higher than that of the object grasped, heat is conducted out of the hand. The faster the rate at which the heat flows, the more rapid the decline in skin temperature, and the more intense the corresponding thermal sensation of coolness. The thermal properties of the object and hand determine the rate at which heat flows. The skin cools more rapidly when in contact with objects made from copper and stainless steel than those made from glass or plastic, and so a stainless steel fork feels cooler than a plastic fork, even though both forks are at the same temperature when picked up. This occurs because stainless steel has a higher thermal conductivity than plastic, which means that when the hand is in contact with the metal fork there is a higher rate of heat flow from the hand to fork.

Thermal cues are not only used to identify and discriminate between objects but can also have a noticeable effect on haptic perception. When people are asked to estimate the heaviness of two weights of equal mass but different temperatures, the colder weight is perceived to be heavier than the weight maintained at skin temperature. A very cold weight of 10 grams (g) is perceptually equivalent to a weight of 100 g maintained at skin temperature. Warming an object also makes it feel heavier than an object at skin temperature, although this effect is smaller and more variable. These effects are presumed to reflect a thermally induced change in the sensitivity of mechanoreceptors in the skin that signal pressure.

Enhancements in tactile acuity thresholds, such as moving two-point discrimination thresholds, have also been observed when the edges and tips of the stimulators in contact with the skin are either cooled or warmed. These perceptual effects are a consequence of the temperature of the object in contact with the skin and not skin temperature. In general, cooling the skin reduces acuity as shown by the decline in sensitivity to changes in pressure, roughness, and vibrotactile stimulation. These latter effects are assumed to result from changes in mechanoreceptor activity as the skin cools.

Lynette A. Jones

See also Cutaneous Perception; Cutaneous Perception: Physiology; Haptics; Multimodal Interactions: Pain–Touch; Pain: Physiological Mechanisms; Tactile Acuity; Texture Perception: Tactile

Further Readings

Dhaka, A., Viswanath, V., & Patapoutian, A. (2006). TRP ion channels and temperature sensation. *Annual Reviews of Neuroscience, 29,* 135–161.

Green, B.G. (2004). Temperature perception and nociception. *Journal of Neurobiology, 61,* 13–29.

Green, B. G., & Akirav, C. (2007). Individual differences in temperature perception: Evidence of common processing of sensation intensity of warmth and cold. *Somatosensory & Motor Research, 24,* 71–84.

Ho, H.-N., & Jones, L. A. (2006). Contribution of thermal cues to material discrimination and localization. *Perception & Psychophysics, 68,* 118–128.

Jones, L. A., & Ho, H.-N. (2008). Warm or cool, large or small? The challenge of thermal displays. *IEEE Transactions on Haptics, 1,* 53–70.

Stevens, J. C. (1991). Thermal sensibility. In M.A. Heller & W. Schiff (Eds.), *The psychology of touch* (pp. 61–90). Hillsdale, NJ: Lawrence Erlbaum.

Stevens, J. C., & Choo, K. K. (1998). Temperature sensitivity of the body surface over the life span. *Somatosensory & Motor Research, 15,* 13–28.

TEXTURE PERCEPTION: TACTILE

Tactile texture perception means the appreciation, by touch, of the characteristics of a surface. In a narrow sense, *texture* refers to the geometry of small irregularities on the surface, such as bumps and ridges (called *texture elements*), but in a broader sense it refers to all the mechanical properties of a surface. Thus, it includes the extent to which the surface gives way (like a ripe pear) when pushed and its frictional resistance (what makes wet hair squeaky) as well. Thermal properties are not included.

A number of perceptual dimensions approximately correspond with these physical properties of a surface. Thus, texture geometry gives rise to the perceptual dimension of roughness/smoothness, resistance to pressure indicates hardness/softness, and frictional resistance gives rise to perceptions of stickiness/slipperiness. We make constant use of these perceptual qualities in everyday life. For example, in walking barefoot across a tile floor, we use perceptions of smoothness, slipperiness, and softness (possibly indicating water) to avoid falling. By far, the most research has been done on the dimension of roughness/smoothness. This entry focuses on the perception of roughness and other dimensions (such as softness and hardness), indirect touch, and virtual surfaces.

Perception of Roughness

Modern work on roughness perception began with a series of studies by Susan Lederman and her colleagues employing special textures called *gratings*. A grating is a texture made of a series of closely spaced ridges, separated by deep grooves. Subjects moved their fingertips across the gratings and gave magnitude estimates of their roughness. Grating roughness was found to depend strongly on the width of the grooves, but to be virtually independent of the width of the ridges. This asymmetry was accounted for by a model in which roughness is proportional to irregularities that are imposed on the skin as the surface presses against it, for example, the amount of skin that bulges into the surface's grooves. Mechanical forces created within the skin by these local deformations are presumably what stimulate receptors and lead to our perceptions. Roughness is only marginally affected by the speed of movement of such surfaces, supporting the idea that spatial, rather than temporal, information is being used to create this aspect of perception.

The physiological basis of roughness judgments was explored by Kenneth Johnson and his colleagues. They recorded the responses of single receptors in the fingerpads of anesthetized monkeys as textured surfaces (arrays of bumps) were moved across them. It was found that roughness was predicted not by the firing rate of any one receptor, but by variations in firing among neighboring receptors. This was especially true in the case of a type of receptor that gives a sustained response to mechanical forces and is therefore called a *slowly adapting mechanoreceptor*. By the time messages from groups of these receptors reach the somatosensory cortex, they combine to influence certain cortical neurons in such a way that the neuron's firing rate reflects the bumpiness of the surface. Activity of these texture-detecting neurons may be the way the brain creates sensations of roughness.

This explanation of roughness perception, in terms of the neural response to a spatial pattern impressed on the skin, accounts for most data on the roughness of surfaces with texture elements larger than about 0.1 millimeter (mm). With smaller texture elements packed closely together, localized skin displacement is negligible, and elements cannot be resolved; yet roughness is still experienced and varies systematically with the geometry (perhaps better called *microgeometry*) of the surface. Clearly a different mechanism for coding roughness must be used for fine textures.

This issue was raised by David Katz as early as 1925. His pioneering experiments suggested that vibrations, set up in the skin as the finger moves across a surface, are involved in our detection of fine textures. Considerable evidence now supports the *duplex theory* that two mechanisms account for roughness perception: The mechanism for coarse surfaces detects spatial characteristics of the surface; the mechanism for fine surfaces detects vibrations that occur as a finger or other sensor moves across the surface.

Evidence for the vibration mechanism is provided by the fact that *vibratory adaptation* (extended stimulation of the skin with vibration, a procedure that is known to temporarily make us less sensitive to it) interferes with the perception of fine, but not coarse, surfaces. The fact that high-frequency vibration (100 hertz [Hz] and above) is especially effective in compromising fine texture perception points to a particular class of mechanoreceptor, the Pacinian corpuscle, as playing a key role. Pacinian corpuscles are large receptors that are specialized for extreme sensitivity to high-frequency vibration. The fact that we are not very good at discriminating fine textures on the face, which lacks Pacinian corpuscles, is further evidence of their importance.

Skin vibrations caused by textured surfaces are complex, like tactile equivalents of rich musical sounds, and vary systematically with the geometrical properties of the texture and the speed of its movement across the skin. Roughness, however, is consistently related to the extent to which a given surface, under a given set of conditions, activates the Pacinian channel.

Many surfaces in everyday life have both a macrotexture and a microtexture: For example, corduroy has both a macrotexture (the wales) and a microtexture (their fuzziness). Work by George Gescheider and colleagues indicates that we are able to attend to either aspect of these complex textures, or, if we choose, to combine both sets of cues into an overall judgment of roughness.

Other Dimensions

Softness/hardness is another perceptual dimension that plays an important role in the overall sensory experience produced by a surface. The research described earlier held this dimension constant by using sets of surfaces all made from the same material. Mandayam Srinivasan and Robert LaMotte examined our ability to experience degrees of softness using sets of surfaces varying in the degree to which they give way when we press on them. It might be thought that kinesthetic cues, such as the extent of forward movement of the pressing finger (detected by receptors in joints), could be used to determine the softness of a surface. However, experiments showed that when the skin is anesthetized, leaving only kinesthetic sensitivity intact, softness judgments are very crude. Thus, it is the stimulation of receptors in the skin that is crucial.

As the finger presses against a surface, the force exerted on the skin increases, whether the surface is hard or soft. But the region of skin in contact with a hard surface remains small, while the finger sinks into a soft surface so that a large skin area is contacted. Pressure (force per unit area) is thus a better indicator of hardness than force is; Srinivasan and LaMotte suggest that we may attend to the rate of pressure increase when judging hardness. Once again, the versatile slowly adapting mechanoreceptors appear to mediate these judgments.

Little is known about how receptors in the skin indicate to the brain that a surface is slippery or sticky. Slip itself is detected when features of the surface (e.g., texture elements) move across the skin. But judgments about friction reflect the relationship between slip and lateral force; kinesthetic cues, or lateral stretching of the skin, may both be involved. Further research is needed to clarify this perceptual process.

Surfaces have values on more than one dimension: A surface might be moderately rough and extremely hard, for example. When stimuli vary along only one dimension in an experiment, it is easy to attend to it alone; but when we examine everyday surfaces that differ from one another in multiple ways, the properties of a surface blend together subjectively into a unified perceptual experience. When people are asked to judge the overall difference between pairs of surfaces, their ratings usually indicate that the surfaces are scattered within a three-dimensional perceptual space, like blueberries in a muffin. Multidimensional scaling (MDS) experiments show that roughness and hardness are two robust dimensions of this *texture space,* and that a somewhat ephemeral third dimension, perhaps slipperiness, is often present as well.

Indirect Touch

We usually touch surfaces directly to examine them, but sometimes another object or material intervenes, as when a doctor examines a patient through a rubber glove, or we feel the texture of writing paper through the tip of a pencil. These are examples of *indirect touch.* Of the two examples given, touching with a rubber glove provides more information, for there is still a distribution of information on the skin, the sort of spatial pattern that texture-detecting cortical neurons of the type described earlier can respond to. Indeed, research in several laboratories has shown that sensitivity to some characteristics of surfaces may be increased by the presence of an intermediate layer of paper or other material that moves with the skin. For example, the material may prevent the microtexture of a test surface from stimulating the skin, thereby making its macrotexture more noticeable.

The type of indirect touch that occurs when we feel the texture of paper when writing with a pencil is, by comparison, severely impoverished, because at any given moment, information from only one point on the surface is reaching the hand. Perception of the surface can still be vivid, but may be altered from the way it would feel with direct touch. For example, if two surfaces with moderately coarse textures are compared, one may feel rougher when they are touched directly, yet the other may seem the rougher one when they are examined through a rigid probe. The reason, as S. J. Lederman and Roberta Klatzky have shown, is that the surfaces that feel the roughest in this type of indirect touch are those whose dimensions just allow the probe tip to drop between texture elements.

Another shortcoming of texture perception through a probe is that, because it depends largely on vibrations conveyed up the stylus, it is very susceptible to vibratory adaptation: Roughness gradually declines when surfaces are repeatedly examined through a probe.

Virtual Surfaces

Despite its shortcomings, indirect touch through a probe has played an important role in the development of virtual surfaces—illusory surfaces created by forces delivered to the hand through a probe. The forces, generated by computer-controlled motors, mimic those that would be produced if the stylus were being drawn across a real surface. The roughness, hardness, and stickiness of virtual surfaces depend on these forces: For example, stickiness is a function of the amount of force that opposes any sideways movement of the probe. There have been reports that a familiar surface can sometimes be recognized if vibrations, recorded when a stylus was drawn across the surface earlier, are "played back" through the stylus to a subject.

In the future, virtual surfaces may be useful in situations where a person is not able to touch a surface of interest directly, because it is far away (e.g., on another planet) or otherwise inaccessible. Sensors in the remote location will detect textural information and relay it to the user's location, where a virtual surface can be created for the user to examine.

This entry has discussed the perception of surface texture without regard to other properties of the examined object, such as its temperature, or other aspects of the situation in which tactile stimulation is occurring, such as sounds produced as the finger moves across a surface. In fact, perceived texture is always just part of our overall sensory experience, and different aspects of that experience have been shown to interact as we work to create a unified understanding of our environment. Continued study of these interactions may provide new insights into the processes underlying texture perception.

Mark Hollins

See also Cutaneous Perception; Cutaneous Perception: Physiology; Haptics; Multimodal Interactions: Pain–Touch; Multimodal Interactions: Tactile–Auditory; Multimodal Interactions: Visual–Haptic; Vibratory Perception; Virtual Reality: Touch/Haptics

Further Readings

Gescheider, G. A., Bolanowski, S. J., Greenfield, T. C., & Brunette, K. (2005). Perception of the tactile texture of raised dot patterns: A multidimensional analysis. *Somatosensory and Motor Research, 22,* 127–140.

Hollins, M., & Bensmaïa, S. J. (2007). The coding of roughness. *Canadian Journal of Experimental Psychology, 61*(3), 184–195.

Hollins, M., Lorenz, F., Seeger, A., & Taylor, R. (2005). Factors contributing to the integration of textural qualities: Evidence from virtual surfaces. *Somatosensory & Motor Research, 22,* 193–206.

Johnson, K. O., & Hsiao, S. S. (1992). Neural mechanisms of tactual form and texture perception. *Annual Review of Neuroscience, 15,* 227–250.

Jones, L. A., & Lederman, S. J. (2006). *Human hand function.* Oxford, UK: Oxford University Press.

Katz, D. (1989). *The world of touch* (L. E. Krueger, Ed. & Trans.). Hillsdale, NJ: Erlbaum. (Original work published 1925)

Klatzky, R. L., Lederman, S. J., Hamilton, C., Grindley, M., & Swendsen, R. H. (2003). Feeling textures through a probe: Effects of probe and surface geometry and exploratory factors. *Perception & Psychophysics, 65,* 613–631.

Srinivasan, M. A., & LaMotte, R. H. (1995). Tactual discrimination of softness. *Journal of Neurophysiology, 73,* 88–101.

TEXTURE PERCEPTION: VISUAL

Our lives are filled with different sorts of texture, and we often make choices based purely on their visual properties. At a glance, we discriminate cotton from burlap from wool, and we can tell how a fabric will feel if we take it between our fingers. We also have no trouble seeing which of two washcloths is moist and choosing the dryer one instead. From the appearance of a wall we can predict what it will feel like when we run our finger across its surface and whether it will leave a mark on our shirt if we lean against it; and when we look across a valley to the face of the mountain on the other side, it is the visual texture projected by that mountain to our eye that tells us what it might be like to try to climb it.

And suppose you climbed that mountain. Imagine picking your way up the trail. Step by step you need to choose the rock most likely to hold your boot sole without slipping. Although this judgment seems natural, it poses a serious computational challenge. It would be hard to build a robot that could make the rapid visual decisions required to hike down a mountain trail. Somehow, your eye and brain need to take in the image of the candidate rock face and quickly decide whether to step on it or not. The image of the rock face is likely to be complex with many semirandom variations in color, and there are likely to be many different sorts of rock faces. How do we make such judgments? Questions like this are central to the field of visual texture perception. This entry describes the history of visual texture perception, models of preattentive texture discrimination, what texture channels exist in human vision, and also revisits texture channels and everyday texture judgments.

History

The instant you open your eyes you see a world made of objects of various shapes and colors. How is this panorama of different qualities created? A natural (but incorrect) theory proposes that all of the differences we see are derived in the retina. We have three classes of cones: the L-cones, M-cones and S-cones, which are sensitive to long, medium, and short wavelength light, respectively. In the 1960s, it had long been known that whenever we see two homogeneous regions as different in color it is because the lights they project to our retina produce different levels of activation in at least one of these three cone classes. Although visual texture is not homogeneous (and therefore produces a range of different activations in any given cone class), it still seemed reasonable to suppose that when we see a difference between two textured regions, it is because one or more cone classes is activated more (on average) by the light projecting to the retina from one of those regions than from the other. In other words, up until the 1960s it was assumed that all of the differences we see spontaneously are sensed by the cones.

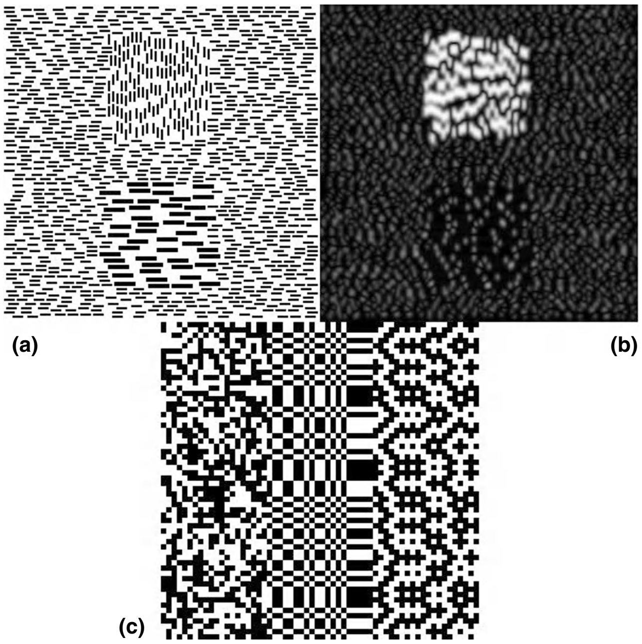


Figure 1 Preattentive Texture Segmentation

Notes: (a) Three preattentively discriminable textures. (b) The response to (a) of a texture channel sensitive to the texture in the upper square. (c) Three isotrigon textures.

That this idea was wrong was demonstrated independently in the late 1960s by Bela Julesz and Jacob Beck, who created images like Figure 1(a) and pointed out that they pose a mystery. This figure contains two square patches of texture set in a background of a third texture. Notice first that each of the two inset patches is immediately evident in this figure. Each patch “pops out” for us effortlessly, proclaiming to our awareness its presence, location, and shape. Such effortless discrimination is called *preattentive* because it happens without any effort on the part of the observer. What makes this display surprising is that the texture in the background and the texture inside each of the two patches all project the same proportions of white and black to your retina. This means that the average spectrum of the light projecting to your retina from within each of the inset patches is identical to the average spectrum projecting from the background. In this case, however, the average response of any cone class must be the same to each of the two texture patches and the background texture in Figure 1(a), implying that the differences we see in this figure are not sensed directly by the cones.

Models of Preattentive Texture Discrimination

How might our vision sense the texture-defined squares in this figure? The fact that these two patches pop out for us automatically (with the same sort of ease as they would if they differed in color from the background) suggests that the strategy used by the visual system to detect them may be similar to the strategy used to detect color-defined squares. Preattentive color discrimination is made possible by the vast sheet of cones covering the back of the eye. This massive array mingles L-cones, M-cones, and S-cones together to simultaneously capture three streams of images reflecting the amounts of long, medium, and short wavelength light projecting to the retina from each point in space. In this retinal array, all cones operate continuously at the same time, enabling the eye to collect vast quantities of image data instant by instant.

Such rapid, image-capturing power is generally assumed to underlie all preattentive discrimination. Thus, the standard model to explain how we see texture-defined figures (like the patches in Figure 1a) proposes that there are a number of arrays of neurons in the visual cortex (in the back of the brain) that operate like retinal cone classes to capture images. However, these arrays are not directly exposed to light (the way the cones are); rather, they receive their input through neural connections to the retina. The images registered by these cortical arrays are assumed to reflect the spatial distribution of other properties than those sensed directly by the three cone classes. Instead of light-spectral properties of the sort sensed by cones, neurons in these cortical arrays are assumed to sense textural properties. For this reason, these arrays are called *texture channels*. Like TV channels, they carry different streams of images, and each stream conveys pictorial information about the changing topography of a particular texture property in the image playing across the retina.

A neuron in a texture channel is assumed to be like a cone in monitoring only a small region of the retina. This region, called the neuron’s receptive field, will generally include a number of different cones and is likely to partially overlap the receptive fields of other nearby neurons in the texture channel array. When an image is projected onto the retina, only the light falling in the neuron’s receptive field has any influence on the neuron’s firing rate. The response of a texture channel neuron reflects a property of the spatial pattern of light falling across its receptive field, and all the neurons in a given texture channel sense the same property but at different spatial locations.

In one of the texture channels, for example, each neuron might fire strongly if the patch of image projecting to its receptive field contains one or more vertical bars of width roughly matched to the width of the bars projecting to your retina from the upper texture patch in Figure 1(a). The only texture in Figure 1(a) that would strongly activate this channel’s neurons is the texture in the upper patch. The neural image produced in this channel by Figure 1(a) would look like Figure 1(b). Each pixel in Figure 1(b) represents a neuron in the texture channel, and the brightness of the pixel gives its firing rate in response to Figure 1(a). Thus, we see that those neurons in this texture channel whose receptive fields are stimulated by the upper texture patch tend to be more highly activated than neurons stimulated by other types of texture.

The texture in the lower patch of Figure 1(a) has the same orientation as the background texture (both textures are made of horizontal bars), but it differs from the background texture in spatial scale (the bars in the lower patch are twice the size of those in the background). The fact that this lower texture patch pops out preattentively from the background texture tells us that at least one of our texture channels is sensitive to the difference in scale between them.

What Texture Channels Exist in Human Vision?

All models that have been offered to explain preattentive texture discrimination include a range of different texture channels like the channel used to generate Figure 1(b) from Figure 1(a). However, these basic texture channels are hypothesized to differ both in their preferred bar orientations and also their preferred bar widths. A particular texture channel would be expected to tolerate moderate variation both in the orientation (with sensitivity dropping to 0 for bars of orientation around 20° above or below the preferred orientation of the texture channel) and also in the spatial scale of the bars it responds to (with sensitivity dropping to 0 for bars double or half the width of the preferred bar width of the channel).

The existence of texture channels of this sort is supported not only by psychophysical experiments, but also by neurophysiology. It has been known since the 1960s (from experiments with monkeys) that the primary visual cortex contains many neurons with receptive fields selective for orientation and spatial scale. In addition, functional magnetic resonance imaging (fMRI) experiments show that neurons with the same orientation and scale sensitivity are arrayed in the brain so that neurons with adjacent receptive fields are adjacent in the brain: exactly the cortical architecture we would expect a texture channel to have. Experiments with monkeys suggest that cortical area V4 is important for some sorts of preattentive texture discrimination; when this region is lesioned, animals previously trained in certain texture discrimination tasks can no longer perform them.

There is also psychophysical evidence for texture channels other than these basic orientation- and scale-selective channels. For example, textures have been constructed that human observers easily discriminate but that cannot be discriminated by orientation- and scale-selective texture channels as they are usually modeled. Three such textures are shown in Figure 1(c). The texture on the left is derived by painting each of its small square elements randomly either black or white; the texture in the middle, called the even texture, is randomized subject to the constraint that the number of black squares in any abutting 2×2 block is even; and the texture on the right, called the odd texture, has an odd number of black squares in any abutting 2×2 block. These three textures are said to be *isotrigon;* this means that it is impossible to tell which of these three types of texture you are looking at if only three squares of the texture are exposed to view (no matter which three squares). It can be proven mathematically that the average response strength of any standard orientation- and scale-selective texture channel will be the same to any two isotrigon textures.

Experiments with textures made of spatially random mixtures of small squares of different gray intensities implicate three intensity-selective texture channels. One channel is sensitive to the average brightness of the texture, another channel is sensitive to the contrast of the texture, and a third "blackshot" channel is sensitive only to the density of very black elements in the texture. Two textures in this class of spatially random mixtures of different intensities are discriminable only if they differ in at least one of these three statistics.

Texture Channels and Everyday Texture Judgments Revisited

So how do you decide which rock to trust your boot to as you hike? Although much work remains to be done to fully answer this question, a reasonable hypothesis says that you combine the neural images carried by your different texture channels, adding them together in exactly the right proportions (which you need to learn through trial and error) to produce a single neural image that reflects (as best as possible) the overall "boot-grippiness" of whichever rock face you're looking at. Then all you have to do is pick the grippiest rock in this image. By combining texture channel images in different ways, you can make a wide variety of different sorts of visual texture judgments, and because of the continuous, image-capturing power of these channels, you can make these judgments very quickly.

Charles Chubb

See also Multimodal Interactions: Spatial Perception in Touch and Vision; Spatial Layout Perception: Neural; Texture Perception: Tactile

Further Readings

Beck, J. (1982). Textural segmentation. In J. Beck (Ed.), *Organization and representation in perception* (pp. 285–317). Hillsdale, NJ: Erlbaum.

Julesz, B. (1975, April). Experiments in the visual perception of texture. *Scientific American,* 34–43.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature, 290,* 91–97.

Julesz, B., Gilbert, E. N., & Victor, J. D. (1978). Visual discrimination of textures with identical third order statistics. *Biological Cybernetics, 31,* 137–149.

Landy, M. S. & Graham, N. (2004) Visual perception of texture. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1106–1118). Cambridge: MIT Press.

THEORETICAL APPROACHES

Theories serve as frameworks to organize facts and as springboards to launch new hypotheses. Some theories in perception are loose collections of assumptions, often unexamined by empirical researchers, and some are well developed and carefully constructed explanatory systems. All empirical research is conducted in the context of some type of theory. The questions asked, the research methodology employed, and the interpretation of the findings are all deeply colored by the theoretical approach taken. Thus, it is important to understand the variety of theoretical approaches that have driven research in perception if one is to be an informed consumer of that research. Many of the facts of perception are familiar features of our everyday experience. Other facts are sometimes surprising laboratory findings. All are dependent on a theory for an explanation of how that fact came to be, and what it tells us about the workings of our perceptual systems.

Consider a sample of facts that a theory must explain: The perceptual world (what we see, hear, smell, taste, and feel) seems to be closely related to what is actually there. Aspects of the three-dimensional (3-D) structure of the world can be known: moving objects appear to move; surfaces appear stable; the perceived attributes of the contents of the visual world (size, shape, color, etc.) remain constant, even as our viewing position changes; the perceptual world is characterized by a high degree of organization and structure; it is a reasonably reliable guide to successful action. Some readers will greet this list with bemusement. Why, they may wonder, is any of this problematic? We see moving objects because objects do move; objects appear to have constant size and shape because these qualities are, in fact, constant. We see, hear, smell, taste, and feel the world as we do because the world is what it is. This stance is sometimes called *naïve realism.*

There is a flaw in the naïve realist position: The naïve realist conflates the facts of the world and the facts of perception. Some qualities of the world are not perceptible to us (the distance to the moon, the freshness of a flower), whereas other qualities may be perceived but inaccurately (the distance to the next mountain, the brightness of one surface in contrast to another). The perceptual world of the honeybee is different from ours. A freshly opened (and thus nectar laden) flower reflects more ultraviolet light than a day old bloom, a beacon to the foraging bee, but undistinguished from its neighbors to us. The qualities of a glass of wine are the same before and after we take a wine-tasting class, but the qualities of our experience surely change. It is how the facts of the world become the facts of perception that is the problem for any theory of perception. Naïve realism has been soundly rejected by the scientific community. The major modern theoretical approaches to the explanation of visual perception are reviewed in this entry, with particular attention to the characterization of the problem and the research questions the approach has motivated.

Picture Theory of Perception

A more sophisticated answer to the question of why the world looks as it does takes into account the optics and function of the eye. Since Descartes peeled away the back of an ox eye and saw the tiny, colorful projection there, theorists have assumed that visual perception is image based. In the *picture theory of perception,* the eye is considered to be a biological camera, designed by evolutionary selection to produce pictures of the world in the form of retinal images. What we can see of the world is limited by what can be contained in the retinal picture. The picture theory of perception is less a theory than a fundamental assumption. The approach is foundational to a number of modern theories, such as constructivism and some versions of the newer computational approach. Images are frequently provided as surrogates for real objects, scenes, and events in laboratory research because differences between the perception of real scenes and the perception of pictures are not expected. Perceptual processes are characterized as processes of image interpretation.

At first glance, the picture theory is attractive, but closer scrutiny reveals serious difficulties. Perceptual theory since Descartes has largely been an effort to explain how we can know what the world is like given the insufficiency of the input (the retinal image) for the job. The image is small, unstable, upside-down, and flat. In addition, it transforms with our movements and is constantly in motion; though for research purposes most have treated it as frozen in time, like a snapshot. Because we can inspect a snapshot and understand something of what is depicted, theorists have supposed that we could have similar success with our retinal images. But any image projected onto a surface such as the retina could have been produced by an infinite number of 3-D configurations in the world. That potted plant apparently growing out of your brother's head in a family snapshot could have been any size and at any distance from him in the world. Pictures are essentially ambiguous, and while photographers, filmmakers, and artists through the ages have figured out ways to make their pictures as meaningful as possible, the fleeting, uncomposed, and unframed images in our eyes have no such advantage. Perceptual theorists have had to look elsewhere for the source of meaning in our perceptual experience. In the 20th century, Gestalt psychologists, constructivists, computationalists, ecological psychologists, and neuropsychologists have all offered solutions. These theoretical approaches are described here.

Structuralism and Gestalt Psychology

Gestalt psychology had its origins in Germany in the early decades of the 20th century in the work of Max Wertheimer, Kurt Koffka, and Wolfgang Köhler. The principal contribution of Gestalt psychology is the claim that the proper object for study in a scientific psychology is everyday perceptual experience. Using the phenomenological method, an observer, confronted with an apple, might report some of the relevant facts of perception, including that the apple is perceived as one object, it is round, red, shiny, and even, perhaps, edible.

Although this claim may seem unexceptional, it was radical in its contemporary theoretical context. *Structuralism,* which dominated experimental psychology in the United States, advocated a mental chemistry that comprised two parts: isolation of the elements or sensations that compose perceptual experience (analogous to a periodical table) and discovery of the laws of integration, which govern the combining of elements (typically laws of association). These aims were to be realized by disciplined application of the method of analytic introspection in which highly trained observers described the putative elementary particles of their perceptual experience (sensations) under tightly controlled conditions. Confronted with an apple, they might report only local values of brightness, hue, and saturation. Notice that in contrast to the Gestalt account, these elements are qualities of visual stimulation and not of the apple itself.

Gestalt theory was further marginalized by the ascendancy of behaviorism after 1920, with its insistence that the only proper object for scientific psychology was overt behavior (and not reports of mental life). Despite this cool reception, Gestalt theory had a significant and enduring effect on the study of perception. The effect is not the result of any specific theoretical formulation related to any particular phenomenon. Instead, the effect was due to widespread acceptance of the Gestalt metatheory (i.e., the claim that everyday perceptual experience defines the products of the perceptual system that need to be explained by perceptual theory: that we perceive the apparent whole, roundness of an apple, for example, even when it is partially hidden by a leaf). The Gestalt psychologists proposed laws of perceptual organization to explain how it is that we see properties of the environment, such as wholeness; and to some extent, the perceptual phenomena identified by the Gestalt psychologists through the application of the phenomenological method still provide the problem set for the evaluation of candidate perceptual theories today.

Constructivism

The cognitive revolution in mid-20th century psychology ushered in a variety of new theories that were not afraid to talk in terms of the operation of the mind. These theories are called "constructivist" because perceptual processes are characterized variously as inferential, logical, intelligent, and creative. There are differences among the varieties of *constructivism,* but they constitute one theoretical approach because they share a set of fundamentals. The starting point for all expressions of constructivism is the unqualified assertion that any effort to develop an account of perception based solely on information in retinal stimulation cannot succeed. The retinal image is simply not rich enough. For example, the perceptual world has three dimensions, while the image has only two. Moreover, as noted earlier, a single retinal image projection is always compatible with an infinite family of configurations in the world. The image is intractably equivocal. On this view, the image is still the starting point for visual perception, but, for retinal stimulation to serve the purpose of supporting adaptive perception, to underwrite visual experience and allow us to get along in the world, the retinal image must be embedded in a context that will enrich the stimulation, disambiguate it, and provide the meaning we need. The constructivist approach proposes two sources of contextual meaning: (1) knowledge of the way the world works that has been instantiated in the visual system over the course of evolutionary history, and (2) knowledge of the relative likelihood of events acquired early in life through interactions with the world.

According to constructivism, the perceptual process is a process of constructing a representation of the world that is simultaneously compatible with these two knowledge bases and retinal input. Theorists couch the process variously in terms of hypothesis testing, logical argument, deduction, and inference. For example, a computer screen obscures a full view of the desk lamp behind it, yet the lamp is perceived as fully intact and present. The retinal image only provides information for the existence of the base and top of the lamp. The rest of the lamp is "constructed" from memories of previous encounters with lamps and knowledge of the way standing objects work. The constructive process is automatic, running its course without conscious control or awareness. For this reason, although the perceptual process is complex, typically perception seems effortless and unproblematic.

At this point what has been presented is not so much a theory as an orientation toward theory, a prescription for the kind of theory that the constructivist would like. To move from the sketch to a full-bodied theory, the constructivist needs to address two questions: What does the perceptual system know (what facts and process rules); and how does the system integrate this knowledge with the retinal input to generate the perceptual world? Ideally, the answer to the second question should have two parts: a model of the perceptual process and a model of the biological hardware that implements the process. Considerable research effort has addressed each of these questions and the approach continues to be fruitful.

Computational Approaches

Computational approaches to perception use computers to model and explore hypothesized psychological and physiological perceptual processes. They often (but not always) take the tenets of constructivism as a starting point. The advent of powerful computers in the last century clearly influenced constructivist approaches to perception. The metaphor of the brain as a computer took firm hold, and constructivist psychological processes were quickly conceived of as algorithms: that is, sets of operations to be performed on input data, using certain rules, to generate an output that is perceptual experience. Not surprisingly, the massive processing power of computers was harnessed in attempts to simulate the ways in which the visual brain was hypothesized to construct the perceptual world. Cameras were hooked up to computers, and the problem of vision was apocryphally assigned to a group of MIT graduate students as a summer project. From the beginning, computational theorists shared the mainstream assumption that biological vision is essentially image based. Early efforts at artificial vision took a still photograph as their input and attempted to find the boundaries, or edges, of 3-D objects by processing the pixels that composed an image of a group of blocks. Getting the artificial vision machine to do anything close to "seeing" the blocks, let alone act on them, proved extraordinarily challenging, revealing more about the impoverished and ambiguous nature of still images and the misguided construal of the process than about biological visual perception. Early machines had trouble distinguishing between edges in the image that belonged to the blocks and edges that belonged to shadows, for example.

Little progress was made in artificial vision until David Marr placed his image analysis algorithms in the context of a theory that emphasized three levels of analysis: a hardware level concerned with the physical structures that perform the task, an algorithm level that sets out the steps that must be taken to solve a problem from input to output, and finally, and most importantly, a computational theory level that articulates the function of the mechanism and the nature of the task to be solved. Marr's approach is a constructivist one. Three-dimensional objects are constructed by extracting successively more information from the image in a series of processing steps, then introducing preexisting knowledge to finally identify the objects.

Marr considered that the principal task of vision was to recognize objects (to identify a cat, a ball, a table, for example). He shared with other theorists the common difficulty that the visual system must identify 3-D objects from impoverished 2-D input. Although Marr's theory more fully described a method for extracting information from the image than his precursors, he, like the constructivists, proposed that the visual system must use prior knowledge of the world to enrich the input and complete the task. Object recognition necessarily involves prior knowledge of the objects to be identified (a soccer ball can be distinguished from a basketball by a knowledgeable observer), and object recognition is certainly an important human achievement, but other theorists have questioned whether it is the central task of visual perception.

The Ecological Approach

A radical alternative to mainstream constructivism, the *ecological approach* claims that the input for visual perception has been misconstrued. Although when light enters the eye images are formed on the retina, the images themselves are not the source of information about the world. Ecological theory rejects the central assumption that visual perception is image based. Instead, it asserts that what is important is that the bundles of light rays reflected from the surfaces of the world before they enter the eye are structured in orderly ways. There are relationships among the light rays: ratios, gradients, and so on, which can be described in terms of angles subtended at the eye by the surfaces of the world. Furthermore, that structure is not ambiguous or impoverished but adequate to specify some properties of the world and events within it. When properly described, the input for visual perception is rich in meaning and particularly well suited to the task of helping us move around in the world and interact with its objects and surfaces. The retina, embedded in a moving eyeball, in a moving head, on a moving body that can interact with the world, is not simply a projection screen, but part of a device that has evolved to sample the structured light, called the optic array. The information about the world contained in this structure is then used to guide behavior. Indeed, the movement of the observer through the world creates flow within the optic array, and this continuous movement is considered to be beneficial and not a source of noise as it reveals the structure of the environment and the relationship of the observer to it as the observer moves about the world. This approach was formulated originally in the writings of J. J. Gibson in a series of three books. Subsequent researchers have been able to demonstrate that the information in optic flow controls the locomotion and actions of many species. It guides honeybees on flights between a foraging site and the hive, in humans it can control finding one's way, steering and braking, ball catching, aperture passing, and more. But there is also evidence that optic flow is not necessary for guiding locomotion. This is a good example of how empirical research has set limits on the generality of a theory.

In contrast to the constructivist approach, prior knowledge of the world is not necessary for perception to proceed, and the process of perception is not characterized as thoughtlike, logical, or inferential. Thus, unlike in the constructivist approach, research does not focus on identifying necessary knowledge or on spelling out specific algorithms. Instead, research motivated by the ecological approach attempts first to identify a particular informative structure in the optic array, and then to demonstrate that the structure is both detected by the perceiver and effective in guiding a particular behavior.

Embodied Perception

Recently, perception researchers have begun to pay attention to the fact that perceptual systems are literally embodied within an observer having certain physical capabilities and behaviors, and that this is relevant to theory building and to the generation of research questions. Furthermore, actions taken by the observer depend upon perception and can themselves create new perceptions, such as when you move your head to discover more about the 3-D structure of an object or test a suspicious surface to check that it is safe to step on. The embodied perception approach challenges two premises that have been widely assumed and, with the notable exception of Gibson and his followers, unexamined by perceptual researchers. One premise is an assumption that the principal operation of the visual system is the generation of representations (descriptions) of the environment. The second premise is that the mind in general and the visual system in particular can be fruitfully understood as a disembodied computer—"a brain in a dish." Lately, both premises have been challenged. Advocates of the ecological theory have presented evidence that the proper objects of perception are not the properties of the environment described by physics (e.g., absolute sizes and shapes), rather, the objects of perception are the possibilities for actions offered by the environment in relation to the observer/actor (so-called *affordances*). We see whether an object is graspable, a gap crossable, a distance walkable. And advocates of embodied perception have presented evidence that perception (say of spatial layout) is significantly influenced by the context of the action in which the observer/actor is engaged. For example, a hill may appear steeper or an object may appear heavier when we are tired or if we are frail.

Another challenge to the traditional view of the disembodied mind is contained in the proposal of David Milner and Melvyn Goodale that there are two visual systems with distinct cortical pathways. Both pathways begin in the primary visual cortex in the occipital lobe of the brain. The ventral pathway extends to the temporal lobe and supports the generation of representations of the world that allow us to think about and report on its objects and events. The dorsal pathway extends to the parietal lobe and supports our ability to act in the world, to reach for and grasp objects.

Empirical research motivated by attention to the embodied nature of perception and by the proposal that there are two visual systems necessarily asks participants to act in the world as well as to report on it. The introduction of action tasks and measures in these new embodied approaches reveals interesting facts about perception that were previously unknown. Embodied perception has thus added an important new dimension to our understanding of perception.

Future Outlook

Modern theories of visual perception differ with respect to their position on several central assumptions: that vision is necessarily image based, that perception generates representations of the world, and that perception can be understood independently of the operation of the body within which it takes place and of the actions it must perform. The research questions raised by the different theories lead to very different styles of experimentation and very different sets of facts. Each approach has contributed a valuable perspective with new insights and important facts. The long-term goal for perceptual theory is to search for a framework that can incorporate all the facts known so far. Such a theory will undoubtedly generate new hypotheses and new facts, and will lead to new interpretations of the facts we have in hand. There may be more assumptions we have not yet thought to question, and new research tasks we have not yet thought to try. Theoretical developments are the engine that drives empirical research and the lens through which we make sense of the facts of perception.

Sheena Rogers and William Epstein

See also Bayesian Approach; Computational Approaches; Direct Perception; Ecological Approach; Embodied Perception; Evolutionary Approach; Evolutionary Approach: Perceptual Adaptations; Gestalt Approach; Indirect Nature of Perception; Philosophical Approaches; Physiological Approach; Psychophysical Approach; Signal Detection Theory and Procedures; Statistical Learning

Further Readings

Epstein, W., & Rogers, S. (Eds.). (1995). *Perception of space and motion: Handbook of perception and cognition* (2nd ed.). San Diego, CA: Academic Press.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Hoffman, D. D. (1998). *Visual intelligence: How we create what we see.* New York: Norton.

Köhler, W. (1947). *Gestalt psychology.* New York: Norton.

Marr, D. (1982). *Vision.* San Francisco: Freeman.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action.* New York: Oxford University Press.

Pastore, N. (1971). *Selective history of theories of visual perception: 1650–1950.* New York: Oxford University Press.

Rock, I. (1983). *The logic of perception.* Cambridge: MIT Press.

THEORY OF MIND

Human adults readily make judgments about the thoughts, perceptions, desires, and intentions behind observable behavior. Imagine that you are at a party. You see Gill with an empty glass, look around, and then approach the drinks table. She picks up a bottle of red wine, makes a brief expression of disgust, pours herself a glass from a bottle of white, and smiles with anticipation. You might suppose that she wanted a drink, likes white wine, but initially mistook the bottle of red for a bottle of white. Such judgments are often derived from directly observable behavior, but allow deeper understanding of what someone is doing, and what they may do next. *Theory of mind* or *mentalizing* refers to this ability go beyond observed behavior and make inferences about unobservable mental states. Most research has examined the development of theory of mind abilities, with broad agreement that children understand concepts such as belief by five or six years of age. However, recent work has suggested that human infants as young as 13 months, as well as some nonhuman species, may pass critical tests of understanding beliefs and knowledge. This, and other evidence from cognitive psychology and cognitive neuroscience, is leading researchers to reappraise what theory of mind is, and how these abilities are linked to other aspects of perception and cognition. This entry explores what is distinctive about theory of mind explanations, how the theory of mind is studied, the cognitive and neural basis of theory of mind, counterevidence, and theoretical accounts.

What Is Distinctive About Theory of Mind Explanations?

We can understand a good deal about everyday behavior just by thinking about observable actions. We see Gill with an empty glass looking around the room and we suppose that she will locate the drinks table and approach it. We see her reject the red wine and take the white wine, and we can predict that she will do the same in the future. But thinking about Gill's beliefs and desires gives additional purchase on the problem of explaining and predicting her behavior. Gill's behavioral history of seeking white wine would not enable us to understand why she initially approached the red wine on this occasion. But if we infer mental states—by supposing that she likes white wine and mistakenly thought this was what was in the first bottle—then we can make sense of her initial selection, her disgust on finding out that she was wrong, and her reselection of another bottle. Such theory of mind explanations gives a more complete account of Gill's actions than explanations that make reference only to observed behavior. However, this comes at the cost of complexity and uncertainty in formulating such explanations.

The key problem is that beliefs, desires, and intentions cannot be directly observed, and must instead be inferred on the basis of observable evidence and background knowledge. Against everyday intuitions, even perceptual states such as "seeing" cannot be observed in a straightforward manner. To illustrate, imagine that you observe a reader and a nonreader viewing a piece of text. Both people show the same observable signs that they are "seeing"—in that both are oriented toward the text and appear to be paying attention to it. However, these observable signs give no notion that the reader sees meaningful words on the page, whereas the nonreader sees meaningless squiggles. Even the apparently simple case of seeing requires us to make inferences that go beyond the observable data in order to understand what might actually be seen. And of course this rapidly increases in complexity if we wish to take account of the many interactions between beliefs, desires, perceptions, and intentions, as we might when evaluating the guilt of a defendant in a court case or the intentions of a lover who is not returning our call. These considerations make theory of mind seem more like reasoning than perception, and this is indeed the prevailing assumption behind many studies of theory of mind. But as we shall see, there are good reasons for supposing that at least some aspects of theory of mind need to be more like perceptual processes than is often supposed.

How Is Theory of Mind Studied?

What would Gill rather do: Stay talking to a bore who is telling her how much they earn or go and find herself another drink? Most people would predict correctly that she would find herself a drink. But do they make this prediction by reasoning about Gill's desires? Perhaps, but we cannot be sure because most people would share Gill's preference and so could make the correct prediction without thinking about Gill's preferences at all. For this reason, a key device in research on theory of mind is to set up a perspective difference between the participant and the target person whose mental states must be contemplated.

To investigate children's understanding of desires, Betty Repacholi and Alison Gopnik presented 18-month-olds, who typically like crackers and not broccoli, with an experimenter who showed a liking for broccoli but not crackers. When the experimenter asked the children for some food to eat, the children passed the experimenter broccoli rather than crackers, clearly choosing according to the experimenter's expressed desire rather than the child's own preference.

By 24 months, and perhaps earlier, children show evidence of understanding that others may see and know about different things—so-called level-1 perspective taking. If an adult shows visual attention toward objects that are out of the child's line of sight, the child may move themselves to a position where they can share the adult's field of view. Conversely, if the child witnesses the hiding of an object but the target adult does not, the child is much more likely to try to communicate the object's location to the adult than if the adult also witnessed the hiding. However, there is no evidence that 24-month-olds understand that the very same scene or object will "look different" to people who view it from different angles—only at around 4 years do children succeed at such level-2 perspective taking.

Four years is also a critical age of change on the most robust and widely studied tasks in theory of mind research: false belief tasks. To assess children's understanding that other people could have different beliefs from them, Heinz Wimmer and Josef Perner created a story in which a character, Maxi, put his chocolate in one location before going outside to play. While he is away, the child witnesses a second character move the chocolate to a new location, so creating a perspective difference between the child and Maxi. The child is then asked where Maxi will first look for his chocolate when he returns. Most 4-year-olds answer correctly that Maxi will look in the old location, successfully taking account of his false belief. However, many 3-year-olds judge incorrectly that Maxi will look for the chocolate in its new location, despite success on questions testing their memory for the sequence of events and the initial and final locations of the chocolate. This finding is extremely robust, having been replicated in over 200 studies using a variety of methods and test questions.

The Cognitive and Neural Basis of Theory of Mind

Evidence from 2- to 5-year-olds suggests that children's growing understanding of desires, perceptions, and beliefs is closely linked with developments in language and executive function (the cognitive control processes necessary for flexible thought and behavior). In adults, it is clear that language is closely integrated with theory of mind, in that much of what we know about other people's mental states comes from what they or other people tell us. However, severe impairment of language following brain injury in adulthood does not preclude success on nonverbal versions of the theory of mind tasks previously described. Executive function continues to contribute to theory of mind abilities throughout the lifespan. Adults' performance on theory of mind tasks adapted from those previously described often correlates with their performance on tests of executive function, and simultaneous performance of theory of mind and executive function tasks results in interference, suggesting that the tasks draw on common processing resources. Moreover, impairment to executive function following brain injury may result in striking difficulties with thinking about other people's desires, beliefs, or visual perspectives, most particularly when those perspectives are strongly at odds with the patient's own. Several studies suggest that inferences about beliefs are neither made automatically, nor necessarily used to interpret ongoing social interaction and communication.

Positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) neuroimaging have shown that a remarkably consistent set of neural regions are involved when adults perform theory of mind tasks, most notably the temporal poles, the medial prefrontal cortex, and the temporoparietal junction/superior temporal sulcus. Converging evidence from studies of brain-injured patients suggests that the temporoparietal junction may, in fact, be necessary for belief reasoning, but the role of the other brain regions is less clear. Studies of developmental conditions such as autism and psychiatric conditions such as schizophrenia suggest that theory of mind may be impaired out of proportion to general impairments in IQ. However, because these conditions may also result in disproportionate impairment to executive function, language, and other cognitive abilities, it remains unclear whether theory of mind is a primary impairment or whether theory of mind impairment is a consequence of impairment to other processes.

There is considerable debate about whether any piece of functional or neural architecture serves a truly specific role in theory of mind. However, the central roles of language and executive function fit well with the view that theory of mind is essentially reasoning with social content, rather than perception.

Counterevidence

If the previous characterization of theory of mind is fully correct, then we should not observe theory of mind abilities in human infants and nonhuman animals, because they are notably lacking in language or executive function. However, recent research suggests otherwise. Several studies of children and infants have adapted false belief scenarios like the one previously described so that looking behavior can be recorded. Whereas 3-year-olds typically predict Maxi's search incorrectly when explicitly asked where he will look for his chocolate, the looking behavior of 2-year-olds anticipates where he will search, and even infants as young as 13 months show surprise (evidenced by longer looking times) if Maxi actually searches correctly in spite of his false belief. Likewise, by making the methods appropriate for the ecology of different species, researchers have recently found evidence that humans are not the only species with theory of mind abilities. Just two examples are that chimpanzees prefer to approach food that a dominant competitor does not know about and avoid food that he or she does know about, and Western scrub jays adapt their food catching behavior depending on whether or not competitors know where food is hidden.

Neuroimaging methods suggest that neural systems involved in first-person action production are also activated when observing someone else's action (consistent with neurophysiologial evidence on mirror-neurons). Importantly, the neural activation is not simply driven by observed action, but by the action that is expected—the perceived *motor intention*—and is also modulated by other cues to the actor's mental states, such as his or her eye gaze. Such evidence has been used to suggest that theory of mind is grounded in these sensorimotor processes.

These findings clearly suggest that infants and some nonhuman species are sensitive to the knowledge states of others, and that action perception is sensitive to social contextual cues. However, whether this is achieved by representing states of knowledge and intention or by representing behavioral rules and expectations is hotly contested. Recall from the earlier example that we can make considerable sense of Gill's actions just by making links between her looking behavior, her approach behavior, and her facial expressions. On a conservative interpretation, the evidence from infants, nonhuman animals, and mirror-neurons might be explained by processing of such observable facts, rather than unobservable mental states. Importantly, whatever it is that infants, nonhuman animals, and mirror-neurons are doing, it is not dependent on language or executive function, suggesting that it is less like social reasoning and more like social perception.

Theoretical Accounts

The tension between theory of mind as social perception versus social reasoning is apparent in the contradictory assumptions of the two dominant classes of theoretical account. In one way of thinking, theory of mind depends on highly specialized knowledge and cognitive processes, constituting a functional module for social perception that does not depend on language or executive function. Such accounts can accommodate the theory of mind abilities of infants and nonhuman animals, and might explain how theory of mind could be fast and efficient enough to sustain ongoing social interaction and communication. In another way of thinking, theory of mind is essentially social reasoning. Such reasoning might use theorylike principles about the role of beliefs, desires, and intentions in mental life, or a process of mental simulation whereby we "put ourselves in someone else's shoes," or possibly both. This second class of accounts can explain how theory of mind is flexible enough for jurisprudence and management of complex social relationships, and would predict a substantial role for executive function and language in theory of mind. These two types of account have traditionally been seen as competitors with mutually exclusive theoretical commitments. However, with the current evidence it seems possible that both may need to be true if we are to explain how theory of mind is at once a task of social reasoning and social perception.

Ian A. Apperly

See also Action and Vision; Causality; Event Perception; Eye Movements During Cognition and Conversation; Infant Perception; Mirror Neurons; Modularity; Social Perception

Further Readings

Apperly, I. A., Samson, D., & Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental Psychology, 45*(1), 190–201.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12*(5), 187–192.

Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology, 50,* 21–45.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron, 50,* 531–543.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences, 8*(9), 396–403.

Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences, 9*(10), 459–462.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology, 55,* 87–124.

TICKLE

See Itch, Tickle, and Tingle

TIMBRE PERCEPTION

Timbre is that characteristic of sound that distinguishes between the perception of two sounds with identical loudness, duration, and pitch. Timbre is often described in terms of the material of the sound source (e.g., woody, metallic), its vocal quality (e.g., nasal, hoarse), its resemblance to musical instruments (e.g., flutelike, bell-like), or its sensory or emotional effect (e.g., sweet, unpleasant). Cross-modal tactile or visual associations can illustrate the timbre of a sound as well (e.g., soft, coarse, dark, bright). Finally, timbre can describe the inferred spatial and reverberant qualities of a sound (e.g., distant or cavernous). This entry covers spectral and temporal attributes of timbre, as well as multidimensionality and exploratory approaches to timbre.

Spectral Attributes

Sounds with discernible pitch generally comprise the sum of multiple sinusoids (i.e., the fundamental tone and its overtones). The frequency of an overtone is, as in many cases including the human voice, an integer multiple of the lowest frequency. The strength of each overtone contributes to the timbre of the whole sound.

Brightness is a well-observed aspect in timbre perception. The vowel "*ee*" sounds brighter than "*oo*," and "glockenspiel" sounds brighter than "marimba" when the sounds are played in the same pitch and loudness. Such brighter sounds often have more energy concentrated at higher overtones, showing greater magnitudes at higher frequencies. Consider, for example, a standard audio equalizer. When the lower frequency range is amplified, music sounds duller, whereas amplifying the higher frequency range causes the same music to sound brighter.

The observation of overtones tells us that resonance plays an important role in timbre perception. When pronouncing vowels, we shape the vocal tract to a form specifically to each vowel, delivering a unique resonance pattern for each vowel. Such resonance patterns can be observed in the relative strengths of the sound's overtones. The overtones around the resonance frequency have stronger magnitude. The distinctiveness between these resonance patterns facilitates the differentiation between vowels regardless of the pitch and loudness. Although some musical instruments share the human voice's ability to produce multiple resonance patterns, many musical instruments produce one characteristic resonance pattern, which contributes to the unique timbre of that instrument.

The timbre of noise is also dependent on resonance. When the high-frequency range of a noise burst is attenuated, the resulting noise is often described as warmer.

Temporal Attributes

The time envelope refers to the overall shape of the waveform of a sound, representing amplitude change over time. For a sound with an onset, steady state, and decay, its time envelope rises, sustains, and decreases. A guitar sound has a quickly decaying envelope, while a trumpet sound can sustain roughly constant strength over the duration of the sound. The rise of a sound may be fast or slow. The time envelope dramatically affects the perception of timbre. For example, a piano tone played backward, reversing its time envelope, sounds rather like an organ tone.

In a typical musical sound, there are many overtones whose respective strengths change over the duration of the sound. This dynamic variation of harmonic content, called spectral flux, also affects timbre perception. For example, the sound of a string instrument often has quickly decaying higher overtones and more sustained lower overtones. Such dynamic properties give richness to the variety of timbre and affect the perceived sound quality.

Multidimensionality and Exploratory Approaches

In the previous sections, the individual factors that affect timbre perception were discussed. However, real sounds feature a combination of all these factors, resulting in a myriad of timbre variants.

When listening or imagining a sequence of many sounds, we notice multiple degrees of timbre similarity: timbres are often categorically perceived in terms of relative similarity or difference. The characteristic timbres of a flute and clarinet (played at relatively moderate loudness and duration) are closer to each other than that of a flute and piano. However, the similarities among timbres are not unidirectional, but multidimensional (unlike pitch and loudness). Researchers have investigated perceptual organization of timbres using methods such as multidimensional scaling (MDS) or semantic differential analysis.

In a MDS study, participants numerically rate the dissimilarity of pairs of instrumental sounds. The statistical analysis (MDS) arranges the sounds in a space with orthogonal axes, given the perceived distances between timbres. According to the

MDS studies, three dimensions (brightness, time envelope, and spectral flux) are known to be predominant.

With semantic differential analysis, participants rate a sound using a set of opposing pairs of adjectives, such as bright/dull or smooth/rough, in the experiment. The statistical analysis (factor analysis) finds the number of dimensions needed to explain the ratings. One study found there to be four dimensions: thick/thin, compact/diffuse, colorful/colorless, and full/empty.

Hiroko Terasawa

See also Acoustics and Concert Halls; Audition; Music Cognition and Perception; Speech Perception

Further Readings

Donnadieu, S. (2007). Mental representation of the timbre of complex sounds. In J. Beauchamp (Ed.), *Analysis, synthesis, and perception of musical sounds* (pp. 272–319). New York: Springer.

Grey, J. (1977). Multidimensional perceptual scaling of musical timbre. *Journal of the Acoustical Society of America, 71*(5), 1270–1277.

Handel, S. (1995). Timbre perception and auditory object identification. In B. Moore (Ed.), *Hearing* (pp. 425–461). New York: Academic Press.

Rossing, T. (2001). *The science of sound* (3rd ed.). San Francisco: Addison Wesley.

TIME COURSE OF PERCEPTION

See Speed of Processing in Sensory Systems

TIME PERCEPTION

At first glance, the idea of perceiving time might seem a bit strange. In the case of visual perception, for instance, there is obviously a stimulus—an object—to be perceived, and a sensory receptor—the eye—for capturing the stimulus. However, it is difficult to imagine a temporal stimulus or even a receptor for time. Nevertheless, situations arise all day long where a person has to take into account his or her impressions about time. For example, if you are waiting at a red light, you will not have to use a watch to realize when it takes an abnormally long time for the light to turn green. Likewise, if someone stops or hesitates in the normal flow of a conversation, the listener will notice the delays—or temporal intervals—very easily. Being aware of the passage of time (i.e., of the duration between events) is what is referred to as *time perception.* This entry describes how experimental psychologists study time perception and present some theoretical viewpoints about the mechanisms leading to judgments about time.

An Internal-Clock Perspective

To explain the ability to make explicit judgments about time, researchers typically refer to a so-called internal clock. The exact nature of this clock is a matter of debate, but a classical view is to assume that there is only one device that takes on all duties related to experiencing time.

A Pacemaker-Counter Device

In many time perception models, researchers assume that the central, internal clock is a kind of pacemaker-counter device. This pacemaker device emits pulses in the form of neural activity in the brain, which is then accumulated in a counter. When a person has to judge the length of a time interval, he or she does so based on the number of accumulated pulses. The errors observed when estimating time are most often attributed to the emitting properties of the pacemaker, and sometimes to the fallibility of the counter.

This simple pacemaker-counter device is the basis of the *scalar expectancy theory.* According to this theory, there is one central, internal timekeeping device that operates for a large range of durations. This theory predicts that the relationship between psychological time and physical time is linear, and that the variability of psychological time (in a series of estimations) is proportional to the magnitude of physical time. This variability to magnitude ratio is constant, which is an application of what is known as Weber's law in perception and psychophysics. In other words, the spreads of the variability for different time magnitudes are superimposed when reported on a common scale (the *scalar* property). Moreover, it is recognized that errors in judging time are not only the result of the inaccuracy of the clock components themselves, but also of the memory and decisional processes involved in a given time-estimation task.

Empirical Approaches

There is much empirical evidence, derived from both animal and human studies, to support the scalar property of time. People participating in experimental studies are sometimes asked to discriminate the relative length of two consecutive time intervals marked by sensory signals. The minimum difference needed to distinguish two intervals, the *difference threshold* as it is referred to in psychophysics, is often reported to increase linearly with the magnitude of the intervals to be discriminated, at least for a certain range of durations. For example, in animal timing literature, it is reported that the scalar property (ratio of threshold to magnitude) holds for a wide range of intervals (from 0.1 to 100 seconds).

As well, participants may be asked to make a finger tap every x millisecond (ms), with two consecutive taps marking one interval length. Instead of discriminating intervals, participants produce intervals. The variability of the inter-tap intervals depends on the magnitude of x. More specifically, for one series of interval productions, this variability (described with the standard deviation—SD) will increase proportionally with the magnitude of the mean (M) intervals produced. Therefore, the SD to M ratio remains constant (another illustration of the scalar property of time).

The Important Role of Attention

It is a common observation that time seems to slow down or speed up depending on whether or not a person is anxiously awaiting an event. For instance, if you are absorbed in an interesting activity, time seems to pass rapidly; if the activity is boring ("a watched pot never boils"), time generally seems to pass slowly. Several empirical investigations, based on the manipulation of attention, have been conducted to explore this avenue systematically. A classical approach is to place participants in a double task condition (i.e., where they are asked to perform both a temporal task, like discriminating or producing intervals, and a nontemporal task, like a mental calculation). Typically, when less attention is allocated to time because of the requirements of the nontemporal task, the duration of an interval is perceived as shorter.

This type of result is consistent with the pacemaker-counter device previously described. Researchers assume that there is a switch, or a gate, that determines the entry of emitted pulses in the counter. This additional process is under the control of an individual's attention mechanism. Therefore, if more attention is allocated to time, more pulses are accumulated and duration is perceived as being longer.

Critical Issues

It makes sense to believe that there is an internal clock, such as a chemical clock, when one considers that judgments about time are influenced by body temperature. For instance, someone with a high fever might have the impression that a minute elapsed after only 40 s. The fact that the scalar property previously described is observed in many experiments reinforces this belief. Additional support is also provided by the fact that levels of performance in different timing tasks, like discrimination and production, are correlated: It is likely that people who do better in one timing task will also do better in another timing task. Nevertheless, there are reasons to question the robustness of this internal clock hypothesis.

First, the time perception literature also shows instances where the scalar property might fail to apply. Consider, for instance, how you would behave if you were asked to estimate a 5.5-s interval. You would probably adopt a segmentation strategy like counting seconds. Although it is not possible to exclude the possibility that keeping track of time with a segmenting strategy might be based on the pacemaker-counter device, the relationship between the variability of estimates and the magnitude of time differs according to whether or not a segmenting strategy is used. With explicit counts, the ratio of variability to time might decrease as time gets longer.

Maybe more crucial is the fact that the accuracy of brief time estimates varies with the nature of the sensory events marking time. For instance, people can estimate an interval marked by two successive brief sensory signals much more precisely when these signals are sounds rather than flashes. This might not be surprising given that auditory functions like speech or music are built basically on the temporal resolution of events delivered in sequences. This modality difference does not negate the existence of an internal clock, as access to the clock might simply be more variable with one type of signal than with another. Nevertheless, it legitimates the perspective that internal time might emerge from properties inherent to each sensory mode.

Alternative Perspectives

Time perception research has been approached in several other ways. Some researchers even argue that there is no time sense, or no direct experience of time, but only temporal effects caused by events. Here are some of the multiple perspectives one may adopt in relation to time perception.

Rhythm and Attending

Suppose you are presented with two intervals marked by three brief consecutive auditory signals and have to decide if the last interval was shorter or longer than the first. Then suppose that six instead of three signals are presented, marking five intervals, the first four being of equal length. If you are asked to decide whether the last interval was shorter or longer than the each of the first four, your judgment will be better than if only two intervals were presented.

This type of finding could be interpreted as resulting from a better, more stable representation in memory allowed by the repetition of the first intervals. Such an explanation is consistent with the clock-process hypothesis previously referred to, as memory processes are recognized to be a potential source of errors when time is estimated. However, this finding could be viewed from another perspective. Repeating a regular sequence of events induces rhythm. Indeed, it makes the arrival of the next event predictable (or coherent). Such coherence, observable in activities such as speech, music, or even locomotion, creates within a person an attending attitude referred to as a *future-oriented attending mode.* According to this view, the capacity to judge time depends on the synchronization of the internal rhythmicity of attending with the appropriate external level of rhythm available in the environment. Hence, the fact that events occur before or after the expected moment of occurrence provides impressions that duration is, respectively, short or long.

Interestingly, this notion of expectation could be applied to much longer intervals (multiminutes long). If you expect that you will have to wait for an event for, say, 15 minutes, and the expected event occurs after only 10 minutes, this 10-minute period will be underestimated (i.e., perceived as having been short).

Biological Issues

Like researchers in most subfields of perception and experimental psychology, those studying time perception have benefited from the recent emergence of neuroscientific investigations. However, linking time perception to specific biological structures is a difficult enterprise. This is because one needs to isolate effects related to some timekeeping process (like an internal clock) from other structures associated with attention and memory processes that are involved in estimating time.

Neuropsychological data have provided evidence that the cerebellum has a computational role in the processing of time. For instance, patients with a cerebellar lesion have more difficulty than those without a lesion in discriminating brief intervals marked with sounds, although the two groups show no difference when it comes to discriminating the intensity of sounds.

Electroencephalography (EEG) has made it possible to reveal the critical role of the frontal regions in the accumulation of temporal information and in the building of the memory trace of an interval. For instance, in a task where a current interval is compared to a pre-memorized target interval, the contingent negative variation (CNV: an expectancy wave) peaked at the end of the current interval (at the end of the accumulation of the so-called time pulses) in the left part of the frontal cortex.

Recent technologies, such as functional magnetic resonance imaging, have been used to localize cerebral activity during temporal processing tasks.

This type of imaging has produced evidence that a structure called the basal ganglia (a group of nuclei deep within the brain), and more specifically the striatum (the input region of the basal ganglia), play a critical role during the encoding of a time interval.

Diversity of Time Perception

The research previously described reflects cases where participants make explicit judgments about time in situations where they know in advance that they will have to judge time (prospective timing or experienced duration). There is also a portion of research, usually dedicated to intervals much longer than a few seconds, where participants are asked to estimate the duration of a task or of an activity only after completing the task or activity. Such an experimental condition is referred to as *retrospective timing* (or remembered duration), and judgments under this condition are known to require the involvement of memory mechanisms. In general, judgments of the same time intervals are about 16% longer and 15% less variable in the prospective than in the retrospective paradigms.

Time perception can also refer to several other pieces of information related to psychological time. For instance, it can be confused with another type of task related to temporal processing, *temporal-order judgments,* which refer to judgments about the order of arrival of two successive sensory signals as opposed to judgments about an interval's duration. Other time perception researchers study the capability of children to learn chronometric units at different scales (seconds, minutes, hours, days, weeks, etc.) or to distinguish notions relative to distance, time, and speed.

Indeed, time perception issues can be found on a large range of the physical time scale. Only those directly related to the psychology of perception were described here. Time remains an elusive but ubiquitous phenomenon; perceiving it is a key component of adapted behaviors.

Simon Grondin

See also Attention: Cognitive Influences; Attention: Selective; Audition: Temporal Factors; Psychophysics: Detection; Sound Reproduction and Perception; Vision: Temporal Factors

Further Readings

Gibbon, J., & Allan, L. G. (Eds.). (1984). *Timing and time perception. Annals of the New York Academy of Sciences: Vol. 423.* New York: New York Academy of Sciences.

Grondin, S. (2001). From physical time to the first and second moments of psychological time. *Psychological Bulletin, 127,* 22–44.

Grondin, S. (Ed.). (2008). *Psychology of time.* Bingley, UK: Emerald.

Jones, M. R., & Boltz, M. G. (1989). Dynamic attending and responses to time. *Psychological Review, 96,* 459–491.

Macar, F., Pouthas, V., & Friedman, W. J. (1992). *Time, action and cognition: Towards bridging the gap.* Dordrecht, Netherlands: Kluwer.

Meck, W. H. (Ed.). (2003). *Functional and neural mechanisms of interval timing.* Boca Raton, FL: CRC Press.

Roeckelein, J. E. (2000). *The concept of time in psychology.* Westport, CT: Greenwood Press.

TINNITUS

Despite being described since ancient times, *tinnitus* remains one of the puzzles of the auditory system. Essentially, tinnitus can be defined as the conscious experience of sound without external cause, and it may be perceived as hissing, ringing, or buzzing. Less commonly, tinnitus can derive from sounds internally generated within the human body, such as otoacoustic emissions, contractions of the middle ear muscles, or the pulsing of blood flow. As described in this entry, the mechanisms underlying this common symptom are complex, and troublesome tinnitus involves both the traditional auditory system and also systems of emotion and reaction within the human brain. Effective therapy to inhibit tinnitus has eluded clinicians, and presently available therapeutic approaches involve counseling and sound therapy to reduce the distress associated with the tinnitus. Recent years have seen a resurgence of scientific interest in tinnitus, resulting in a substantial body of research, and this is an area where progress is eagerly anticipated.

Many people experience tinnitus without distress: In the United Kingdom, transient spontaneous tinnitus has been reported in a third of the adult population. Troublesome tinnitus is perceived by 5% of the adult population, and in 0.5% the symptom is debilitating. It has been estimated by the American Tinnitus Association that 12 million Americans have sought medical advice about tinnitus, and 2 million Americans are significantly debilitated by the experience of tinnitus. In such severe cases, tinnitus may be associated with insomnia, poor concentration, irritability, and emotional distress, and unsurprisingly there may be issues of anxiety or depression. Background neuronal activity similar to tinnitus is almost ubiquitous: Experiments where people without tinnitus listen intently in an anechoic chamber demonstrate that over 80% of people can perceive some tinnitus-like activity in such circumstances.

Underlying Mechanisms

In some people, the development of tinnitus is an indication of significant otological pathology, such as an acoustic neuroma, Ménière's disease, or otosclerosis. As such, an interested and informed clinical opinion should be sought by a person with troublesome tinnitus, especially when it is unilateral or pulsatile. In other cases, there is no demonstrable pathology, and the mechanisms of the ignition of the tinnitus are complex. In some cases, the person becomes aware of the background auditory neuronal activity as previously described. In other cases, a hearing loss (usually of cochlear origin) is associated with the tinnitus, partly as the reduction in external auditory information increases the perceived intensity and clarity of the internal tinnitus signal. Physiological theories of tinnitus generation abound, and this reflects the great variety of experience seen in the clinical population of people with troublesome tinnitus. Examples of ignition sites for tinnitus include cochlear hair cell damage following noise-induced hearing loss, disinhibition of central auditory pathways following reduced input to the dorsal cochlear nucleus due to hearing loss, and aberrant plastic reorganization of higher auditory centers following cochlear dysfunction.

Describing these physiological mechanisms of tinnitus does not adequately address the clinical picture of severe, debilitating tinnitus. The consensus view is that the mammalian auditory system has a protective and defensive function, and is vigilant for sounds that might be associated with intrusion or threat, irrespective of intensity. Once such a signal is perceived, systems of reaction and arousal become active, and the animal experiences an orientating response to the signal, to which it pays intent attention. These responses are instinctive and extremely rapid, and are evident in human beings that hear a possible intruder at night or a change in breathing of a newborn child and are instantly alert and vigilant. Once a sound is identified as background or harmless, the animal habituates to that stimulus.

In the development of troublesome tinnitus, the tinnitus signal is perceived as an intrusive or threatening stimulus, evoking arousal and alertness, and further attention to and prioritization of the tinnitus. Over time, the tinnitus may then come to evoke emotional states of apprehension or fear, further reinforcing awareness of the tinnitus. Habituation fails to occur due to the associated emotional distress and heightened levels of agitation and arousal. This understanding of tinnitus distress was synthesized into a "neurophysiological model of tinnitus" by Pawel Jastreboff, an auditory neuroscientist, and Jonathan Hazell, an otologist, in the 1990s, though the component concepts were being widely utilized. Further, the insistence by these authors that the development of troublesome tinnitus must involve classical conditioning has attracted criticism from some in the psychological community who believe this view to be simplistic. Pure psychological models of tinnitus have also been criticized for their potential overemphasis upon emotion. The consensus view at present is that troublesome tinnitus should be viewed from both physiological and psychological perspectives.

Treatment

In the absence of effective pharmacological therapy for tinnitus, traditional treatment has been reassurance, though reports of this being dismissive or clumsy abound. Tinnitus retraining therapy is a treatment paradigm deriving from the work of Jastreboff and Hazell and involves intensive counseling about the nature of tinnitus and mechanisms of distress, plus sound therapy to reduce the starkness of the tinnitus utilizing wide-band sound generators or (less commonly) hearing aids. These devices are worn during waking hours; and environmental sound generators emitting the sounds of rain, ocean waves, and the wind may be utilized at night. The aim is to induce habituation to the tinnitus thereby reducing the associated distress, and in some cases, the perceived awareness of tinnitus. An alternative therapeutic approach is to utilize cognitive behavior therapy techniques to change beliefs (cognitions) about tinnitus and to address the emotional and reactive responses to the tinnitus signal. Comparing the relative efficacy of these treatments is not straightforward, as both involve counseling and so cannot be examined using standard randomized placebo controlled trials. Observational trials indicate that both approaches significantly reduce the impact of tinnitus in 70% of cases. Elements of both styles of treatment are in common use to treat tinnitus around the world, and it is rare for the pure form of either to be utilized.

Future Outlook

It is widely accepted that this situation regarding tinnitus treatment is not adequate, and many scientists and clinicians are working both to increase knowledge regarding mechanisms of tinnitus and to improve the efficacy of treatment. Consideration of pharmacological treatment for tinnitus is ongoing, and many different compounds are under investigation, including selective serotonin reuptake inhibitors commonly used for the treatment of depression. Stimulation of auditory centers within the brain, using electricity or magnetism, is an area of considerable interest, with initial results showing promise. What is evident is that this neglected area is now receiving increasing attention and research resources, and as such, progress is anticipated.

David Baguley

See also Attention and Emotion; Audition; Cortical Reorganization Following Damage; Hearing Aids

Further Readings

American Tinnitus Association: http://www.ata.org

Andersson, G., Baguley, D. M., McKenna, L., & McFerran, D. (2005). *Tinnitus: A multidisciplinary approach*. London: Whurr.

Jastreboff, P. J., Hazell, J. W. P. (2004). *Tinnitus retraining therapy*. Cambridge, UK: Cambridge University Press.

Snow, J. B. (2004). *Tinnitus: theory and management.* Hamilton, Ontario, Canada: BC Decker.

TOOL USE

A tool can be defined as any object that provides a specific mechanical advantage when carrying out a task. *Tool use,* then, can simply be thought of as the act of using an object to help achieve a specific goal. Although the range of tools used by humans is extensive, this entry focuses on the use of common hand tools, such as hammers or spatulas.

Both humans and some other species are known to use tools, but the sophistication of their tool use differs considerably. It was once thought that only humans were capable of tool use, but in recent decades, tool use has been observed among a variety of nonhuman species. For example, among the natural behavior of wild chimpanzees, the use of sticks to collect termites and rocks to crack open nuts easily qualifies as tool use. Surprisingly, recent research has also found that even nonprimate species, such as the common crow, can build and use tools. Among early hominids, stone tools first clearly appeared over two million years ago with the genus Homo, specifically *Homo habilis,* meaning “handy man.” The extent and complexity of this early hominid tool use is believed to closely mirror that of modern chimpanzees. Eventually, however, there was an explosive jump in the complexity of tool use among early humans, accompanied by large increases in cranial capacity. This exponential expansion ultimately enabled the human species to develop ever more sophisticated tools, such as computers or satellites, that go far beyond mere mechanical advantages.

The cognitive abilities required to recognize and use tools are considerable. Consider the example of a key. A key is defined more by its utility than by its visual characteristics. Though visually dissimilar, a modern car key, an antique skeleton key, and a tubular key for a bicycle lock share a common function and pattern of associated actions. The action sequence—picking up a key with the finger and thumb at the base, inserting it in a lock, and turning it—is so automatic that it requires little thought. Tools can also be used in novel ways. For example, when trying to cut open a sealed box without any other tools, a key may be used to tear the tape. This requires an understanding of the desired tool's mechanical properties, such as sharpness, rigidity, and elongation, to enable it to fit beneath and cut the tape. This example illustrates three main cognitive components of tool use: knowing "what" a tool is, knowing "how" to use a tool, and understanding "why" a tool affords a particular mechanical advantage.

Problems Using and Knowing About Tools in Patients With Brain Damage

Interestingly, these three components—"what," "how," and "why"—can be dissociated within the human brain. Some patients with brain damage know "how" to use a tool despite an inability to name "what" the tool is. That is, these patients' problems are restricted to recognizing, naming, and knowing about the functions of tools. Surprisingly, despite such deficits, at least some of these patients can still use tools appropriately. Other patients show the opposite pattern in which they know "what" but not "how." They can recognize and describe features of a given tool, such as its name, function, and context, but they simply cannot carry out the correct movements to use it. Finally, for a third type of patient, the fundamental problem is not in the "what" or "how," but rather in knowing "why" a tool is appropriate for accomplishing a particular task. These patients are often confused about the goals of a particular action and fail to appreciate how a given tool can provide an advantage in fulfilling those goals. For example, a patient with these kinds of deficits may attempt to brush her teeth with a comb or he may try to eat soup with a butter knife. When faced with a problem, like when trying to drive a nail into a piece of wood, these patients fail to see how a hammer would be helpful.

Most of what we know about the brain areas important for tool use comes from the study of patients with brain damage and problems with tools, like those previously described. All of these deficits are most pronounced following damage to the left hemisphere of the cerebral cortex. The most consistent finding is that when patients have problems with the "how" aspects of tool use, they have damage to parietal and/or frontal areas, particularly the left inferior parietal cortex. Less is known about the brain areas important for the "what" and "why" components of tool use; however, areas within the temporal cortex are likely to be critical. Over the past decade, advances in functional neuroimaging techniques, like functional magnetic resonance imaging, have led to additional insights about the brain areas important for tools. The most reliable activations are found within the left posterior temporal gyrus, the left posterior parietal cortex, and the left inferior frontal gyrus.

Tools as an Extension of the Body

Tools may even become an extension of the body. In studies with macaque monkeys, Atsushi Iriki and colleagues have noted remarkable changes in neurons (brain cells) within the parietal cortex as monkeys learn to use a rakelike tool to gather food. Before monkeys were trained to use the rake, these neurons responded only to visual stimuli near to the hand. However, as monkeys became experts with the rake, these same neurons now responded to the range of space accessible with the rake. One interpretation is that these neurons, which code reachable space, can be expanded to include a tool, as if the arm itself had extended. Recent research has also shown that neurons in the frontal cortex of the macaque brain respond when the monkey is using a pair of pliers to grasp an object, suggesting that these neurons encode the goal of an action, not the means by which it is accomplished.

Jody C. Culham and Kenneth F. Valyear

See also Action and Vision; Brain Imaging; Embodied Perception; Reaching and Grasping; Visually Guided Actions

Further Readings

Johnson-Frey, S. H. (2004). The neural bases of complex tool use in humans. *Trends in Cognitive Sciences, 8*(2), 71–78.

Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences, 8*(2), 79–86.

Rothi, L. J. G., & Heilman, K. M. (1997). *Apraxia: The neuropsychology of action.* East Sussex, UK: Psychology Press.

TOP-DOWN AND BOTTOM-UP PROCESSING

Top-down and bottom-up processing refers to the integration of information from one's own cognitive system (top-down) and from the world (bottom-up) to facilitate perception. Imagine that you are preparing for a dinner party, and that you have put a homemade pie in the oven. Some time later, you are in the adjacent dining room, setting the table. The timer goes off and you return to the kitchen to take out the pie. The question posed in the current entry is the following: Upon re-entering the kitchen, how did you "know" which object was the oven? Recognizing the oven may have seemed effortless; however, this ease belies the fact that such recognition results from a complex integration of different processes that each rely on different types of information. Perception comprises an integration of top-down and bottom-up information, with their relative importance determined as a function of the available sensory information, the task at hand, the perceiver's goals, and past experiences.

Bottom-Up Processing

Bottom-up processing is the processing of information from the world that falls on your sensory receptors. With respect to vision, such information might include features corresponding to the color, size, and shape of the oven. With respect to audition, such information might include the frequency and loudness of the buzzer on the oven timer that indicated that the pie was done. With respect to olfaction, such information might include the odors emanating from the now-cooked pie. Note that such information resides in the world, external to you. You only begin to perceive it when the information stimulates your receptors. Thus, the visual information was present and available, but not perceived by you when you were in the dining room; this is because you couldn't see the stove. In contrast, the auditory and olfactory information were perceived, as their features were able to stimulate your receptors even when you were in the adjacent room. Bottom-up processing of available information as it stimulates your sensory receptors constitutes the beginning of perception.

Top-Down Processing

Top-down processing is processing that operates on information that is internally generated from within your own cognitive system, formed by learning and experience. One type of top-down information is your general knowledge of objects, including what they typically look like (in terms of size, shape, color), what they are used for, how they function, and where they are typically located. For example, you know what ovens typically look like, and know that they usually have certain parts (a door, a heating unit) that are in certain relations. One can think of this as a template for an oven. You may use these templates to help you recognize the oven when you go into the kitchen by comparing candidate kitchen objects (oven, refrigerator, sink) to this template. When the input matches this template, you can be said to recognize the oven. Note, however, that these templates need to allow for some variety—for example, they need to take into account the fact that some ovens will also have a cooking surface, whereas others will be built into the wall. This variability can be accommodated with multiple templates all tied to the concept oven or with a single template with flexible features.

A second type of top-down information is your specific knowledge of the objects—not just your understanding of ovens in general, but the specific information about your oven. This personal information is more detailed, including which features your oven has—as opposed to which features it could have. For example, you know that your oven is self-cleaning, that it was a special-order purchase, and that it has a light switch in the top-left corner, but that the light no longer works. The difference between these two types of top-down knowledge can be illustrated with an example. You have the knowledge that generally kitchen appliances (including ovens) tend to be white or black—these are the colors on the models that are displayed in home improvement centers; however, you also have the knowledge based on your own specific experience that your oven is olive green, a retro-color that you selected when you redecorated your kitchen last year.

Interplay Between Top-Down and Bottom-Up Processing

Sometimes the top-down and bottom-up information compete. In the oven color example, upon entering the kitchen, your general knowledge of ovens might lead you to expect to see a black or white appliance with typical size and shape; however, this expectation would be contradicted by the green color information that was stimulating your visual receptors. The reconciliation of this conflict would be based on your knowledge that you purchased an atypically colored oven, and that in this case the bottom-up information in combination with the specific knowledge about your prior purchase should be prioritized over the more general top-down expectation.

In addition, the relative strengths of each of these types of information may vary, depending upon the specific situation. On one hand, if you have a lot of top-down expectation, you may rely less on the bottom-up information, and simply encode less of the input from the world—only enough to verify that it is consistent with your expectations. For example, when you rush into the kitchen to take out the pie, you may use size, shape, and location to help recognize the oven, but you may not notice changes, for example, the replacement of a faded green kitchen towel on the handle of the oven with a new brightly colored red towel. On the other hand, when there is not a strong top-down expectation or not a lot of experience with a given object, you will need to rely more on bottom-up information. For example, upon entering your home, a guest would need to rely on bottom-up information about the aroma in the kitchen to identify that a pie was baking.

Recognition of ovens seems a trivial example; however, our interactions with the world are full of such seamless and rapid acts of recognition that integrate top-down and bottom-up processing. Another example would be looking for your parked car in a crowded parking lot, in which you combine your knowledge of the color and make of your car with your current visual perception of the cars in the parking lot. You would also use your memory of where you parked to direct your attention to certain areas of the parking lot, restricting your bottom-up processing to places where you expect your car to be. A final example from the tactile modality would be of a Braille reader using his or her knowledge of language (letter and word sequences) to interpret a series of raised dots on a page as meaningful text.

With respect to the relative importance of these types of information, one might expect that the bottom-up sensory information to always have priority—after all, that is the information that is currently stimulating the senses. However, there is some compelling research showing that all of the sensory information that stimulates the receptors may not necessarily be processed. Indeed, we seem to be remarkably poor at detecting changes to objects and scenes when they occur during a discontinuity, such as an eye blink, an eye movement in ongoing perception, or during a change in scene or a cut in a movie. This phenomenon is referred to as *change blindness*. For example, if one is viewing a movie in which some features of the objects in the scene change across cuts (for example, the protagonist's shirt changes color, a lamp on a table disappears), the viewer may not necessarily detect such changes. This failure to detect changes occurs not only in laboratory tasks in which researchers construct movies to intentionally include such changes in order to determine what can and will be detected, but also in commercially produced movies in which these changes are unintentionally made by the filmmaker. This failure to detect changes across discontinuities is mediated somewhat by whether one is intentionally looking for changes, and by whether the changed object is one that had been attended. Nevertheless, even under optimal conditions, it is clear that perception does not operate by recording all visual aspects of the world. If perception did work this way, then one would be able to simply remember all details before the discontinuity and compare them to all the details after the discontinuity, and thereby detect a change if it had occurred. Rather, and counterintuitively, such failures to detect changes suggest that we encoded and represent only very little bottom-up detail from the world. That is, rather than making sure that we store and retain all features of all objects, we assume that these features will stay constant unless there is a reason for them to change. For example, upon rushing into the kitchen to take out the pie, you may fail to detect that a person in the household doing laundry has replaced the green towel that was hanging on the oven handle with a fresh red towel. Rather than encoding and storing in memory the initial color of the towel, you instead simply assume that the color of the towel will remain constant, and that if at any future point in time you want to know its color, you can simply look at it and use bottom-up information to provide you with the color. Kevin O'Regan has characterized this idea as using the external world as a memory store. Note, though, that when you do look at the towel and see that it is red, you may or may not realize that this is a new towel that has replaced the old green towel; you may think that the red towel was there all along. This is consistent with assumptions that we make about an environment being unchanging, and using the current details to fill in things that we may have previously failed to encode.

Another example showing the interplay between bottom-up and top-down information comes from research by Philippe Schyns and colleagues showing that the features that are used during recognition are not a fixed set but are constructed in a dynamic manner based on task constraints and object information. This is referred to as *diagnostic recognition.* For example, an object may have a large number of possible features that could be used to assist recognition. In the example of finding your car in the parking lot, you may use its color, its make, your memory of its location, a distinctive bumper sticker, and so on. Depending upon the context, you may prioritize certain features over others. For example, if there are numerous red cars in the same area as your car, searching on the basis of color will not help you find your car as much as searching for your distinctive bumper sticker. Then again, depending upon the direction from which you are approaching the car, you may not be in a position to see the bumper sticker. In that case, using your memory of its location may be the most reliable, and then you may need to inspect each car individually.

The interplay between top-down and bottom-up processing is not restricted to recognition in the visual domain. Consider a paradigm case from the auditory domain: speech perception. There are at least two sources of difficulty. First, conversation often takes place in a noisy environment in which there are numerous distracting sounds (other conversations down the hall, the hum of the heating system, the pinging of your e-mail alert). Thus, the cognitive system must somehow filter the myriad sounds impinging on the receptors in order to selectively attend to the sounds that correspond to the speech stream. This has been referred to as *auditory scene analysis.* Second, even when you can functionally isolate the speech stream, the speech signal is usually a fairly continuous sequence of sounds, without breaks that would indicate where one word ends and another begins. Nevertheless, by using the top-down knowledge of individual sounds and knowledge of the words and the structure of sentences in the language, a listener can parse the stream into sequences of meaningful words. In contrast, an unfamiliar foreign language usually sounds like a long continuous sequence of sounds, and it is difficult to identify any individual words. The bottom-up information in both cases is the same. However, when you listen to a familiar language, you also have top-down knowledge that you can use to segment the stream into meaningful words and phrases; when you listen to a foreign language, such top-down knowledge is missing.

Laura A. Carlson

See also Attention: Cognitive Influences; Bayesian Approach; Context Effects in Perception; Speech Perception; Vision: Cognitive Influences

Further Readings

Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Sciences, 8,* 465–471.

Humphreys, G. W., Riddoch, M. J., & Price, C. J. (1997). Top-down processes in object identification: Evidence from experimental psychology, neuropsychology and functional anatomy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 352,* 1275–1282.

O'Regan, K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology, 46,* 461–488.

Schyns, P. G. (1998). Diagnostic recognition: Task constraints, object information, and their interactions. *Cognition, 67,* 147–179.

Simons, D. J. (Guest Ed.). (2000). Change blindness and visual memory. *Visual Cognition*, *7*(1–3).

TRANSCRANIAL MAGNETIC STIMULATION

Today, toys, cars, cookers, TVs, and hi-fi equipment are built with computer technology that means every minor fault seems to necessitate buying a replacement or paying an expert a sum close to the cost of buying a replacement for repairs. Before this, we all had a good chance of working out how things worked because they could be taken apart, tinkered with, repaired, and rebuilt. Many brain scientists belong to this taking apart, tinkering, and fixing tradition. Some investigate the effects of lesions in animals and accidental brain damage in humans and in doing so they can ask which parts are necessary for specific functions. Over the past 25 years, it has become possible to interfere with human brain functions safely and reversibly, and to control when and where the interference is induced. The technique that allows brain scientists to do this is known as *transcranial magnetic stimulation* (TMS). This entry describes inducing currents in the brain, modifying brain excitability, and specificity of TMS in space and time.

Inducing Currents in the Brain

TMS works on the principle of electromagnetic induction discovered by Michael Faraday. An insulated coil carrying a large, brief current is placed over the skull that overlies the part of the brain one wishes to stimulate. The current needs to be brief because induction partly depends on the rate of change of the electromagnetic field. The electrical current creates a brief magnetic field that passes through the scalp, and the magnetic field then induces an electrical current in the brain.

This simple method has many different uses. After over a century of effort the problem of using magnetic fields to stimulate the brain was solved by Anthony Barker in 1985, and it has since been used widely to stimulate both peripheral nerves and brain tissue in studies encompassing motor conduction in human development, motor control, movement disorders, swallowing, vision, memory, speech and language, epilepsy, depression, stroke, pain, and plasticity. It has proved to be a versatile technique and is now also being used in combination with electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and single unit electrode recording in animals.

If the coil carrying the current is placed over the motor cortex, a region of the brain that controls movement, then the subject's finger, arm, hand, face, or leg may twitch because the magnetic field has induced a current in the corresponding region of the brain's body map. If the coil is placed at the back of the subject's head, that is, over a region controlling vision, the subject may see points of light or moving shadows. These are known as phosphenes. One use of TMS, then, is to map the organization of the body in the brain or to map visual space in the brain.

Modifying Brain Excitability

There are a few parameters under the experimenter's control, and varying them leads to TMS having varied uses beyond mapping. Pulses of TMS can be applied at different intensities; they can be applied singly or in repetitive transcranial magnetic stimulation (rTMS) of low or high frequency. The choice of stimulation parameters determines whether the effects of stimulation are excitatory or inhibitory. For example, two single pulses separated by less than 5 milliseconds (ms) can produce intracortical inhibition, that is, they make the stimulated area less sensitive to stimulation. Two single pulses separated by a gap greater than 10 and less then 30 ms can produce intracortical facilitation, that is, make the stimulated area more sensitive to stimulation. Repetitive TMS at a frequency of 1 hertz (Hz) has the effect of depressing cortical excitability for a period of time after the train of pulses has finished, whereas repetitive stimulation at 10 Hz or more may increase excitability. Theta burst stimulation, applying trains of 50 Hz stimulation in bursts every 200 milliseconds, has the effect of depressing cortical activity for a period following stimulation.

Specificity of TMS in Space and Time

The most common type of coil used is a figure of eight shape so that the effects of the current add up where the wires meet in the center of the coil. This has the effect of giving TMS a degree of focality such that specific brain regions can be targeted for stimulation. Although stimulation effects are maximal in the brain region directly underneath the coil, TMS also has secondary effects on connected areas of the cortex, and these are useful in both basic and applied studies. If one stimulates, say, the left motor cortex that controls movement of your right hand, there are three likely effects of stimulation: a change in activity in the targeted region; a change in activity in immediately surrounding areas of the cortex (perhaps the area controlling your right forearm); and a change in the activity of cortical areas directly connected with the stimulated region (most likely to be the adjacent area in opposite side of the brain, so in this case, the region of the brain that controls movement of the left hand). These are important considerations in preventing naïve interpretations of the effects of TMS and also in allowing for studies of how different brain areas interact.

Brain interactions can be studied using TMS because the pulses are brief (approx 1 ms), and by applying a TMS pulse at two different places a few thousandths of a second apart, one can ask how activity in one area affects activation in a second area. One can also ask how long it takes for information to travel between two different brain areas. This temporal specificity of TMS is perhaps its most important contribution to understanding brain function. A great deal of what we know about the organization of the human brain comes from studies of patients with brain damage, but the patients are often very slow in their behavior.

One of the most important contributions of TMS to understanding human brain function is that it can test whether a brain region is necessary for a task. This is especially important in the current era of fMRI, which shows correlations between brain activity and behavior. If one stimulates an area that is activated in a brain imaging study while the subject performs the same or a similar task as in the imaging study, then, if the brain activation records an indispensable function, the TMS will interfere with the function. If, however, the brain activation was a correlate of some inessential component or something about which the brain imager was unaware, the TMS will not have an effect on the behavior.

Methods related to TMS place two electrodes on a person's head and pass a constant or pulsed current (in the region of 1–3 millampere [mA]) between them in the hope that some of the current will pass through the scalp and reach the brain. This is called transcranial direct current stimulation (TDCS) if the current is constant and transcranial alternating current stimulation (TACS) if the current is pulsed.

Vincent Walsh

See also Brain Imaging; Cortical Organization; Modularity; Neural Recording

Further Readings

Walsh, V., & Pascual-Leone, A. (2003). *Transcranial magnetic stimulation: A neurochronometrics of mind.* Cambridge: MIT Press.

Wassermann, E., Epstein, C. M., Ziemann, U., Walsh, V., Paus, T., & Lisanby, S. (Eds.). (2008). *The Oxford handbook of transcranial stimulation.* New York: Oxford University Press.

TUNING CURVES

The neurons within each sensory system must collectively encode the vast ranges of stimulus information that are environmentally relevant to the organism. A ubiquitous characteristic of neurons in all sensory systems, however, is their sensitivity to a wide range of stimulus parameters. That is, no single neuron is so selective that it is sensitive to just one single stimulus value. For example, in the auditory system, there will be no neuron that responds only to, say, middle C (~262 hertz [Hz]) and no other sound frequencies. In order for each of the sensory systems to encode the ranges of possible stimulus parameters, the neurons comprising those systems divvy up the range such that each neuron exhibits differing selectivity, but over restricted ranges of those stimuli. For example, in the auditory system, the responses of neurons are often jointly sensitive to sound frequency and intensity, with different neurons encoding different ranges of frequency and intensity. The range of frequencies over which an auditory neuron is responsive is indicated by its *tuning curve*. In the visual system, peripheral neurons are differentially selective for the spatial location, luminance, and wavelength of the stimulus. This entry examines the concept of the tuning curve, which is an experimental measure of the selectivity of neurons in sensory systems to one or more stimulus parameters.

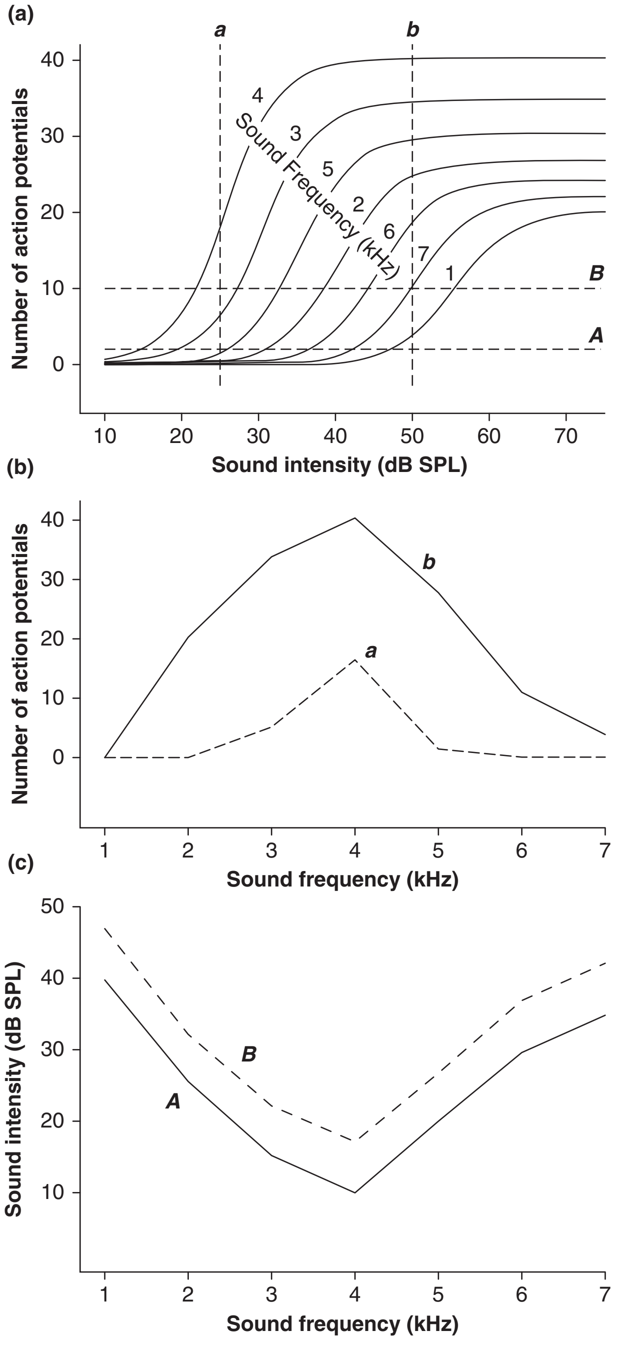


Figure 1 Example of a Tuning Curve

Notes: (a) The experimentally measured response of an auditory neuron in terms of the number of sound-evoked action potentials is determined as a joint function of two different stimulus parameters, frequency (kilohertz [kHz]) and intensity (sound pressure level in decibels [dB SPL]). (b) Tuning curve for sound frequency given a fixed value of intensity. Tuning curves are shown for two values of sound intensity in panel a (corresponding to dashed lines *a* and *b*). (c) Tuning curves that result from two different constant criterion values of the number of action potentials in panel (a) (corresponding to lines *A* and *B*).

Measuring Tuning Curves

Figure 1 shows an example of how tuning curves are typically measured experimentally. Tuning curves can be measured as a function of one stimulus parameter or jointly as a function of two stimulus parameters. Here is a concrete example from the auditory system where tuning curves are often determined by measuring the number of action potentials a neuron fires (ordinate of Figure 1a) while varying the frequency (parameter of Figure 1a) and the intensity (abscissa of Figure 1a) of sounds presented to the ears. In many cases, the "responses" (action potentials in the current example) of the sensory systems being studied are related monotonically to monotonic increases in at least one stimulus parameter. This is particularly true for neurons or receptors at the most peripheral levels of a system.

It is important to keep in mind that quality and the quantity of the "response" being measured experimentally in the generation of a tuning curve can be of many different types. Most common of these is the number of action potentials a neuron under study fires in response to the presentation of a stimulus, such as in the example in Figure 1. The intracellular voltage (or current) potential of peripheral sensory receptors (e.g., hair cells, photoreceptors) and central neurons can also be used. These are just a few common types of responses that are measured in the generation of tuning curves.

For the purposes of this entry, the number of action potentials in response to sounds will be used to designate the magnitude of the experimentally measured response to variations in one or two of the stimulus parameters. These parameters in this example are sound frequency (kHz) and intensity (sound pressure level in decibels, dB SPL). Moreover, the physical magnitudes of frequency and intensity vary monotonically in a physically meaningful way; in the example frequency is changed from 1 to 7 kHz and intensity from 10 to 70 dB SPL. This latter issue is important in order to interpret the functional meaning of the tuning curve; if the stimuli do not vary in a physically meaningful way, it is not a tuning curve. As in the example in Figure 1, in the auditory system, dB SPL and frequency can be varied in a monotonic fashion that is physically meaningful in that both SPL and frequency can be related directly to the physical characteristics of sound energy.

Figure 1(a) shows the number of action potentials as a joint function of sound intensity and frequency. Sound intensity is plotted monotonically increasing along the abscissa, whereas sound frequency is the parameter of the graph and is indicated next to each respective curve. Each of the curves in Figure 1(a) corresponds to the number of action potentials that results with a fixed value for sound frequency while varying the sound intensity. For example, the left-most curve results from fixing the frequency of the sound at 4 kHz and then varying sound intensity from 10 to 70 dB SPL. Two general attributes of sensory responses are apparent in Figure 1(a). First, as one of the stimulus parameters (i.e., intensity) is incremented monotonically, there is a concomitant monotonic increase in the number of action potentials, regardless of the sound frequency. In general, this finding will hold true only over limited ranges of both parameters. The empirical determination of what these stimulus ranges are is one of the primary purposes of measuring a tuning curve! Second, for a fixed value of sound intensity, there is a *non-monotonic* change in the number of action potentials as the sound frequency is varied. Indeed, it is this non-monotonic change in the neural response that exemplifies the concept of a tuning curve.

Figure 1(a) simply illustrates how the neural responses change in response to joint variations in two stimulus parameters. But plotting the data in this way obscures the tuning curve. In general, there are two ways in which tuning curves are often measured experimentally and plotted in the literature. Figure 1(b) shows one of the ways. Here, the number of action potentials is plotted as a function of monotonic variations of just one of the stimulus parameters, in this case sound frequency, while the other parameter (intensity) is held at a constant value. In the example in Figure 1(b), the tuning curve for frequency is plotted for two arbitrary values of intensity. The specific values for the intensity are indicated by the vertical dashed lines designated by *a* and *b* in Figure 1(a). The tuning curve for frequency, given intensity, is indicated by the systematic change in the number of action potentials evoked from the neuron being studied. Note that there is a particular value of frequency for which the number of action potentials is maximal (i.e., 4 kHz), and the response decreases for values of frequency both above and below this “best” value. In audition, the stimulus frequency yielding the maximal neural response is often called the best frequency. Again, the response of the nervous system can be of a variety of different types, depending upon what is being studied. The generic tuning curve illustrated in Figure 1(b) is arguably the most common method by which the selectivity, or tuning, for stimulus parameters is measured empirically in sensory systems. In addition to the auditory example, some additional examples from other sensory systems include photoreceptors tuned to particular ranges of the wavelength of light; neurons in the primary visual cortex tuned to the orientation of bars of light; mechanoreceptors in the somatosensory system tuned to the frequency of mechanical stimulation; and cold and warm thermoreceptors tuned to cool and warm temperatures, respectively.

Figure 1(c) illustrates the other common method of measuring a tuning curve, primarily in the auditory system. For this type, the tuning is characterized with respect to a fixed, or criterion magnitude of the number of evoked action potentials. That is, the values of the sound frequency and intensity presented to the ear are jointly adjusted so as to maintain a constant number of evoked action potentials. Figure 1(c) shows the tuning curves that result for two different constant numbers of action potentials, corresponding to the horizontal dashed lines *A* and *B*. Conceptually, this kind of tuning curve is measured by holding constant the value of one parameter, say frequency, and then systematically incrementing the other parameter, intensity, until the desired number of action potentials is evoked. Then sound frequency is incremented and the process of finding the value of sound intensity to achieve the same number of action potentials is repeated again. This process is repeated for each desired value of frequency. The resulting tuning curve illustrates the minimum value of one stimulus parameter, intensity, given the other parameter, frequency, that is just required to achieve the desired number of action potentials. This type of tuning curve is ubiquitous in studies of the auditory system. The criterion level is typically, but not always, chosen to be the minimal neural response (e.g., number of action potentials) that is just significantly greater than the "noise" in the neural responses being measured. Of course, this "noise" level can mean a variety of things depending upon what is being measured experimentally: For receptor potential this might be the intrinsic noisiness of the intracellular voltage, whereas for action potentials from central neurons, this would be the spontaneous rate of action potentials from the neuron under study. The spontaneous response is the intrinsic response of the neuron that occurs in the absence of any external stimulus.

For the type of tuning curve shown in Figure 1(c), several important pieces of information can be measured. First, the minimum, or trough, in the tuning curve indicates the sound intensity to which the system is most sensitive. This intensity value is called the *threshold* if the criterion neural response (e.g., number of action potentials) is set to be near the smallest measurable response (e.g., the spontaneous rate of a neuron). The frequency at which the threshold occurs indicates the frequency about which the system is most selective. In the auditory system, this is called the *characteristic frequency* because this frequency is characteristic of the place along the cochlea from which this frequency selectivity arises. Finally, a quantitative measure of the frequency selectivity of the system can be determined from the width of the tuning curve; that is, given a sound intensity over what range of frequency does the system still respond.

Interpreting Tuning Curves

Tuning curves can be used in a variety of ways to study sensory systems. For example, for a given sensory system, tuning curves for some stimulus parameter (or combination of parameters) are often measured at different anatomical levels along the ascending pathways, from periphery to the central nervous system (CNS). This kind of systematic measurement can give experimenters important clues about how the neural representations of stimuli are being transformed from one level of the nervous system to another. The degree of selectivity for different stimuli as measured by a tuning curve may give an indication of the acuity of a neuron or how well the neuron is able to resolve or separate stimuli. This has important implications for behavior, as it is often believed that the behavioral ability to discriminate differences in sensory stimuli is limited by the selectivity of the peripheral neurons in that system to those stimuli. For example, auditory neurons with sharp frequency selectivity (i.e., responsive over a narrow range of frequency, as in Figure 1c) may be necessary for the behavioral ability to discriminate two sounds that have only slight differences in their frequencies.

Daniel J. Tollin

See also Audition; Audition: Pitch Perception; Auditory Frequency Analysis: Neural; Auditory Frequency Selectivity; Color Perception; Neural Representation/Coding; Physiological Approach; Sound Stimulus; Vision

Further Readings

Bitterman, Y., Mukamel, R., Malach, R., Fried, I., & Nelken, I. (2008). Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. *Nature, 451,* 197–201.

Bolanowski, S. J., Jr., & Verrillo, R. T. (1982). Temperature and criterion effects in a somatosensory subsystem: A neurophysiological and psychophysical study, *Journal of Neurophysiology, 48,* 837–856.

Galambos, R., & Davis, H. (1943). The response of single auditory-nerve fibers to acoustic stimulation, *Journal of Neurophysiology, 6,* 39–57.

Henry, G. H., Dreher, B., & Bishop, P. O. (1974). Orientation specificity of cells in cat striate cortex. *Journal of Neurophysiology, 37,* 1394–1409.

Kiang, N. Y.-S., Watanabe, T., Thomas, E. C., & Clark, L. F. (1965). *Discharge patterns of single fibers in the cat's auditory nerve* (M.I.T. Research Monograph No. 35). Cambridge: MIT Press.

Russell, I. J., & Sellick, P. M. (1978). Intracellular studies of hair cells in the mammalian cochlea. *Journal of Physiology, 284,* 261–290.

Schnapf, J. L., Kraft, T. W., & Baylor, D. A. (1987). Spectral sensitivity of human cone photoreceptors. *Nature, 325,* 439–441.

U

Unconscious Processes

The ease with which we open our eyes and consciously experience the world around us belies the complexity of the underlying neural mechanisms, which remain incompletely understood. Many of these neural mechanisms operate outside of awareness. For example, our conscious perception of an object in the immediate environment, such as this book, must depend on neural processing that acts on the pattern of light incident on the retina and ultimately leads to accurate identification of that pattern of light as representing this book. Yet we have no conscious access to those computations, only to the outcome of the process that somehow results in awareness of an object in the environment. Neural activity must therefore reflect both conscious processes (the outcome of object identification) and unconscious processes leading to object identification.

Although conscious and unconscious sensory processing have quantitative and qualitatively different effects on brain activity and behavior, the accurate delineation of the boundary between conscious and unconscious perception based on behavioral reports remains challenging. However, an increasing amount of evidence suggests pervasive unconscious processing of stimuli in the human sensory cortex. This entry describes unconscious inference, behavioral measures of awareness, qualitative differences between conscious and unconscious perception, neural correlates of unconscious processing, and unconscious processing after brain damage.

Unconscious Inference

One influential and persistent notion in the study of perception is the idea that it represents or depends upon unconscious inference. This phrase is conventionally attributed to the 19th-century physiologist Hermann von Helmholtz, who proposed that our percepts are not the direct products of sensory stimulation alone. Helmholtz argued that sensory stimulation is so impoverished and unconstrained that it cannot account for the richness and apparent precision of our conscious experience. Instead, the phenomenal character of visual awareness must arise from a series of unconscious reasoning processes applied to sensory input. These processes consist of a series of inferences based on premises about knowledge of the world acquired through experience. Although Helmholtz provided the first explicit conceptual formulation, the underlying ideas date back to some of the earliest accounts of vision. For example, Greek philosophers such as Euclid believed that the apparent size of an object reflected the visual angle it subtended at the eye. But if we hold our thumb up to the night sky, we do not perceive our thumb to be the same size as the moon even if they subtend the same visual angle at the eye. Five centuries after Euclid, the Greek philosopher Ptolemy therefore proposed that apparent size depended both on the visual angle an object subtended at the eye and its perceived distance. How big something appears therefore depends both on sensory stimulation and on inferences about how far away it appears to be. Perception therefore depends on the combination of sensory input with acquired knowledge in an inferential process that does not reach awareness.

Helmholtz believed that unconscious inferences underlying perception were similar in character to those that underlie our ability to perform general or scientific reasoning. The notion that our ability to engage in rational argument and unconscious perceptual inference relies on common cognitive machinery associated with general reasoning or intellect has subsequently come under sustained attack. In particular, the mechanisms underlying sensory perception appear to be more encapsulated, domain-specific, and inference-like rather than general. Such a narrower account of unconscious inference can explain how visual illusions (failures of inference) can persist even when the observer has knowledge of the nature of the illusion. Nevertheless, the notion that perception reflects an ability of the perceiver to combine information provided by the environment with knowledge acquired through experience remains a powerfully articulated central influence on contemporary discussion of perception.

Behavioral Measures of Awareness

Around the same time as Helmholtz, the emerging discipline of psychophysics sought to determine whether it was possible to quantify a threshold for sensory stimulation to result in conscious experience. This threshold could be absolute, specifying the minimum stimulus intensity required to give rise to awareness of the presence of that stimulus, or relative, specifying the smallest change in intensity of a stimulus that would give rise to awareness of a change. The ability to establish such thresholds for sensory stimuli raises the possibility of delineating the boundaries between conscious and unconscious processing empirically through behavioral reports.

Psychophysics relies on observers directly reporting their experience. Critically, this depends not only on the sensitivity of an observer's brain to a visual stimulus, but also on his or her decision criterion (an internal state) for reporting it. This has the consequence that such reports can be unreliable as a measure of whether a sensory stimulus has reached awareness or not. For example, if observers are uncertain about their conscious experience, they may adopt a conservative decision criterion and report uncertain or unclear perception of a stimulus as a failure to perceive that stimulus. This can lead to errors in describing the relationship between perception and brain activity or psychological processes. For example, if an observer erroneously reports unclear perception as absent perception, the underlying psychological processes (or brain activity) may be incorrectly characterized as unconscious.

Such concerns have led to the development of objective methods that factor out the effects of confidence and decision criterion on measures of awareness, such as signal detection theory. Because these furnish estimates of whether a sensory stimulus is consciously discriminable independently of the criterion adopted by an observer, they circumvent the inferential problem previously described. But such approaches require relatively large numbers of trials to determine sensitivity accurately, and if sensitivity varies over time, the possibility of trial-to-trial fluctuations in awareness cannot be ruled out.

Recently, alternate methods for judging whether a stimulus has reached awareness have been proposed, such as postdecision wagering. This approach asks observers to make a sensory discrimination and subsequently place a financial wager on the outcome of that discrimination. A failure to maximize earnings from the wagering even following correct discriminations is taken to indicate that observers are unaware of the outcome of the discrimination (otherwise they could have placed bets to accurately maximize their earnings).

All of these methods have been used in an attempt to show that unconscious perceptual processes exist and can affect behavior. Such a claim remains contentious for some because it is always possible to question such behavioral measures on the grounds either that they are determining whether awareness exists some time after a stimulus has been presented (thus raising the possibility of fleeting awareness not subsequently remembered) or that they are not measuring all the relevant aspects of conscious experience.

However, there now exists considerable evidence consistent with the notion that sensory stimuli that do not reach awareness nevertheless can be processed and influence behavior. This indicates that unconscious processing plays an important role in perception. For example, visual masking studies show that a sensory stimulus rapidly followed by a mask can influence subsequent judgments on related stimuli, demonstrating the existence of unconscious priming. More recently, the advent of noninvasive technologies for functional brain imaging has led to an explosion of interest in the neural correlates of such unconscious processes that are examined in this entry.

Qualitative Differences Between Conscious and Unconscious Perception

Rather than attempting to prove the existence of unconscious perceptual processes, a complementary approach proposed by some authors is to instead *assume* that there is a meaningful conceptual distinction between conscious and unconscious perception, and now inquire whether conscious and unconscious perception differ in their qualitative effects on behavior. To some extent, such an approach sidesteps questions of whether the boundaries between conscious and unconscious processing can be determined from behavioral reports by instead focusing on qualitative differences in behavior. For example, it is now well established that consciously and unconsciously perceived words can have different influences on behavior in a word completion task. In such experiments, observers are briefly presented with single words, followed by a "word stem" comprising the first three letters of that word. They are then asked to complete this word stem with the first word that comes to mind apart from the word that has just been presented. When the words are presented very briefly (so that they are unconsciously perceived), then observers have great difficulty following the instruction and often complete the word stem with the previous word. In contrast, when the words are presented for a longer duration and reach awareness, then the observers are good at following the instructions. This example of how unconsciously and consciously perceived words can have qualitatively different effects on behavior supports the idea that unconscious processes tend to be automatic and uncontrolled, whereas conscious experience is required to exert control and guidance of actions in accord with instructions.

Neural Correlates of Unconscious Processing

The ability to characterize the behavioral boundary between conscious and unconscious processing opens up the possibility of relating such reports to brain activity. Experimental paradigms have been developed that measure changes in neural activity associated with changes in a stimulus in the absence of any changes in conscious experience. This dissociates activity associated with conscious experience from activity associated with unconscious perception.

This type of experimental paradigm has now been employed in a significant number of brain imaging studies. Taken together, they show that many regions of the human sensory cortex can be activated by sensory stimuli that do not reach awareness. The best-studied area of perception to date is vision. Activity can be identified in the absence of visual awareness at the earliest retinotopic cortical stages of visual processing and can be specific both to the location in the visual field and the elementary visual attributes of the stimulus (such as orientation).

This activation of the visual cortex by stimuli that cannot be accurately reported continues to higher stages of processing. For example, activation of functionally specialized areas is consistently observed for masked words, faces, and objects that observers do not report seeing. Unconscious activation also extends to cortical areas considered part of the dorsal stream of visual processing; images of tools presented to just one eye and rendered invisible by a highly salient pattern being flashed in the other eye (intraocular suppression) nevertheless activate the human dorsal stream. Such neuroimaging evidence converges with electrophysiological studies showing that electrical potentials associated with higher levels of visual processing can be elicited by invisible stimuli. Finally, subcortical pathways can also show activation in response to subjectively invisible and unreported emotional visual stimuli. Many of these neuroimaging studies rely on subjective behavioral measures of awareness and have been criticized for not eliminating the possibility of low-confidence conscious perception confounding trials where masked stimuli were not reported, as previously discussed. Nevertheless they show striking convergence with the rather small number of experiments that assess awareness employing objective behavioral measures also previously discussed.

Often, these studies of unconscious neural processing have used masking or interocular suppression approaches to render stimuli invisible. Neural signals elicited by such masked or suppressed stimuli are generally substantially lower than for the equivalent unmasked stimuli that reach awareness. Such observations have led to theoretical claims that conscious perception of a given visual attribute resides in the extrastriate area specialized for that attribute (e.g., area MT/V5 for motion). However, several studies using paradigms other than masking show that invisible stimuli can nevertheless elicit neural activity equivalent in amplitude to visible stimuli. The degree to which signals associated with invisible stimuli are attenuated therefore appears to depend on the experimental paradigm rather than visibility per se, though this merits further investigation.

Sensory stimuli can therefore elicit considerable processing outside awareness, and this can influence subsequent behavior. Why might this be so? The neurobiologists Francis Crick and Christof Koch have argued that the brain may consist of a series of specialized unconscious "zombie" systems that can control behavior on the basis of sensory signals but in the absence of awareness. Such zombie systems raise the question of why the brain does not simply consist of a series of such specialized systems. Crick and Koch argue that such an arrangement would be inefficient in circumstances where many such systems are required, such as for complex organisms capable of generating multiple behaviors such as humans. In such situations, the processing of unconscious systems might be better used to produce a single more complex representation that is available to make a choice among different possible plans for action.

A complementary research strategy to asking which brain structures show activity correlated with unconscious processing is to establish whether differences exist in the pattern and character of neural activity associated with stimuli that reach awareness compared to those that do not. So far, such studies have largely focused on visual perception. They show that visual awareness is associated with changes not only in the level and timing of neural activity, but also in the precise cortical areas that show such changes in activity. For example, a common but not invariable finding is that awareness of a particular visual feature or object in the environment is associated with enhancement of neural signals. This enhancement can be identified in both the primary visual cortex and functionally specialized regions of the higher visual cortex whose neuronal specificities represent that stimulus attribute or object category.

In addition to the amplitude of neural activity, the precise timing of that activity can also matter. In some experimental situations, conscious and unconscious processing can be dissociated only at a relatively late stage following presentation of a stimulus, after several hundred milliseconds. This has led to theoretical suggestions that unconscious processes are represented by a "feed-forward" sweep of information processing with conscious perception correlating only with later feedback or recurrent signals.

Finally, although the visual cortex plays a central role in representing the contents of visual consciousness, there is now considerable evidence that activity in the frontal and parietal cortex is strongly correlated with changes in the contents of visual awareness. For example, activity during transitions in binocular rivalry, and other forms of bistable perception, is time locked to frontal and parietal cortex activity. This type of evidence suggests not only that unconscious processes might be associated with certain types of neural activity, but also that such processes might be restricted to the primary sensory cortex and association areas. However, such a conclusion is at best tentative and remains under active empirical investigation.

Unconscious Processing After Brain Damage

Striking evidence for unconscious processes can also be observed following brain damage associated with stroke or brain tumors. For example, damage to the primary visual cortex is typically associated with a lack of awareness for stimuli presented at corresponding points in the visual field. However, when patients with such cortical damage are asked to guess properties of stimuli they deny be being able to see when they are presented in a scotoma, several can show residual visual capacity in their blind field. These patients are able to perform certain discriminations and localizations better than chance in the acknowledged absence of awareness. This ability has become known as *blindsight* and suggests an important role for V1 in visual awareness.

Damage to the frontal and parietal cortex is commonly associated with visual neglect and visual extinction, where patients do not perceive or respond to any type of visual stimulus placed in one half of the visual field. Nevertheless, stimuli presented in their neglected visual field can still influence their behavior. For example, when asked to choose between a picture of an intact and a burning house that differ only in the part of the image placed in the neglected visual field, patients with parietal neglect reliably choose the intact house. Such abilities must depend on unconscious processing of stimuli placed in the neglected visual field. Consistent with this, neuroimaging reveals that extinguished and unseen stimuli placed in the left visual field of a patient with parietal neglect still give rise to activity in corresponding regions of the contralateral visual cortex. Moreover, this unconscious processing of visual stimuli extends to encompass category-selective processing, with unconscious activation of face-selective regions of the ventral visual cortex when a face is presented compared to nonfacial stimuli. These data show that activation of visual cortices is insufficient to result in awareness following parietal damage and that unconscious processes proceed as far as identification of the category of a visual stimulus despite absent awareness.

Geraint Rees

See also Attention: Covert; Attention and Consciousness; Change Detection; Consciousness; Consciousness: Disorders; Extrasensory Perception; Neuropsychology of Perception; Priming

Further Readings

Cheeseman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology 40,* 353–367.

Crick, F., & Koch, C. (2000). The unconscious homunculus. In T. Metzinger (Ed.), *The neuronal correlates of consciousness* (pp. 103–110). Cambridge: MIT Press.

De Gelder, B., de Haan, E. H. F., & Heywood, C. A. (Eds.). (2001). *Out of mind: Varieties of unconscious processes.* Oxford, UK: Oxford University Press.

Frith, C., Perry, R., & Lumer, E. (1999). The neural correlates of conscious experience: An experimental framework. *Trends in Cognitive Sciences, 3*(3), 105–114.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 362,* 857–875.

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences, 23*(11), 571–579.

Weiskrantz, L. (1998). *Blindsight: A case study and its implications.* Oxford, UK: Oxford University Press. (Original work published 1986)

V

VESTIBULAR SYSTEM

A common experience of childhood is spinning in circles and then stopping and feeling the world and oneself to be turning. These after-sensations are the result of stimulation of receptors in the *vestibular system,* or labyrinth, of the inner ear that are differentially specialized for the detection of angular and linear acceleration. On either side of the head, there are three semicircular canals embedded in the temporal bones, orthogonally oriented in relation to one another, and also two otolith organs. These are the receptor mechanisms that detect angular and linear acceleration, respectively. They are illustrated in Figure 1. When a head movement is made in or near the plane of rotation of a semicircular canal, the endolymph fluid within the canal lags and "displaces" a door-like membrane structure, the cupula, in the canal that is embedded with the cilia of receptor cells. Over the frequency range of natural head movements, because of the visco-elastic properties of the cupula and the endolymph in the semicircular canal, the neural output of a canal is actually proportional to head velocity in space. When a person voluntarily spins, the cupula of each canal in the plane of rotation is initially displaced and then, during constant velocity rotation, returns to its rest position. When the person stops, the endolymph fluid in the canals lags and displaces the cupulae in the opposite direction. As a consequence, a sense of rotation is experienced even though the person is stationary. This entry covers the control of eye movements, posture, and orientation, motion sickness, and models of vestibular function.

Control of Eye Movements

Semicircular canal signals are important in the control of eye movements. They provide effective ocular stabilization for displacements of the head. For example, when a gaze shift is made in terms of a combined eye and head movement to look at a target located off to the side, the eyes will initially move and reach the target and then the head will start to move. As the head moves, the semicircular canals are activated and with very short latency they innervate eye muscles that counter-rotate the eyes in the direction opposite the movement of the head. In this way, stable fixation of the target is maintained during a shift in direction of both the eyes and head. Without this compensatory stabilizing mechanism, the movement of the head following the movement of the eyes would carry the eyes past the target. These so-called vestibulo-ocular reflexes (VORs) are quite exact for pitch (up-down) and yaw (left-right) movements of the head.

There are a variety of disorders involving the semicircular canals. For example, occasionally an otoconium composed of calcium carbonate crystals from the otolith organs can break loose from the otolith membrane and become lodged against the cupula of a semicircular canal. This creates an unusual situation because, normally, the functioning of a semicircular canal is independent of

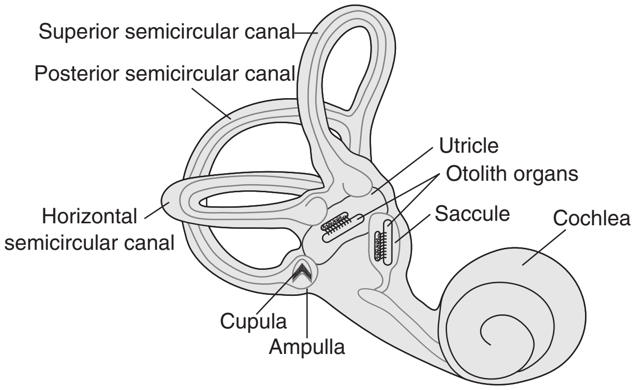


Figure 1 Arrangement of the Three Semicircular Canals and the Utricle and Saccule

the direction of gravity. The endolymph in the semicircular canal and the cupula, the little door within the canal, have the same density. As a consequence, the canals are not affected by static head orientation in relation to gravity because there is no turning couple related to gravity as the fluid and cupula are homogeneous in density. By contrast, when an otoconium with a specific gravity of 2.7 lodges against the cupula (which has a density of about 1.005), it renders the canal gravity dependent. Consequently, when the canal is oriented so that the force of gravity pushes the otoconium against the cupula, it displaces it and activates the receptors of the canal. In this circumstance, the affected individual will feel as if he or she is turning even when stationary. This quickly can lead to disorientation, loss of balance, and motion sickness.

The bilaterally symmetric semicircular canals are organized in a push-pull arrangement so that when one is excited its mate on the other side is inhibited. The canals also have very high resting discharge rates. This turns out to be an advantage because if one canal is damaged, after a period of adaptation its partner canal will take over and be able to signal rotation of the head in both directions. A resetting of a central mechanism takes place so that for the remaining canal an increase in discharge above rest levels will be associated with compensatory eye movements in one direction and a decrease in discharge will drive eye movements in the opposite direction.

Control of Posture and Orientation

The other receptor system of the labyrinth, the otolith organs, responds to linear acceleration including the force of gravity. There are two bilaterally symmetric otolith organs on each side of the head, a utricle and a saccule. When a person is standing with head tilted forward about 30º, the utricle membrane is horizontal and the saccule membrane is oriented vertically, as shown in Figure 2. Gravity is a linear form of acceleration acting on the body, and, consequently, for each orientation of the head relative to gravity a shear force is generated on the otolith membranes. The otoconia embedded in the membranes have a density much higher (~2.7) than that of the surrounding tissues (~1.05) and thus displace the membranes as a function of head tilt. Cilia embedded in the membranes are bent and the extent of their bending is coded in the discharge rate of the axons of their parent cell bodies. These signals determine the head tilt relative to the upright position that will be experienced.

Linear accelerations applied to the body also stimulate the otolith organs and this presents a potential problem, one that has arisen with the development of vehicles and aircraft that can impose substantial linear accelerations on the

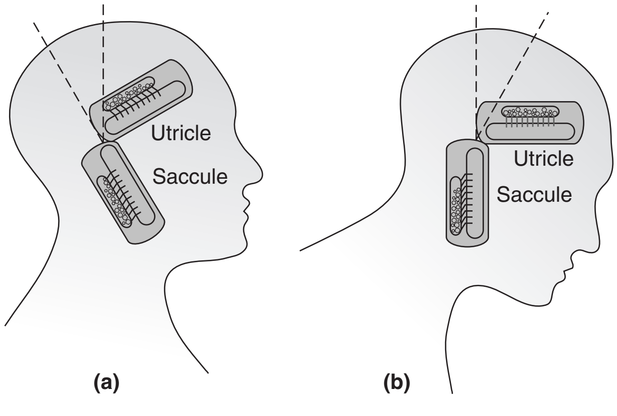


Figure 2 Head and Utricle Membrane Angles

Notes: This figure shows that when the head is upright (a), the utricle membrane is tilted about 30º upward creating a shear force on the membrane, and when the head is tilted 30º forward (b), the embrane is horizontal.

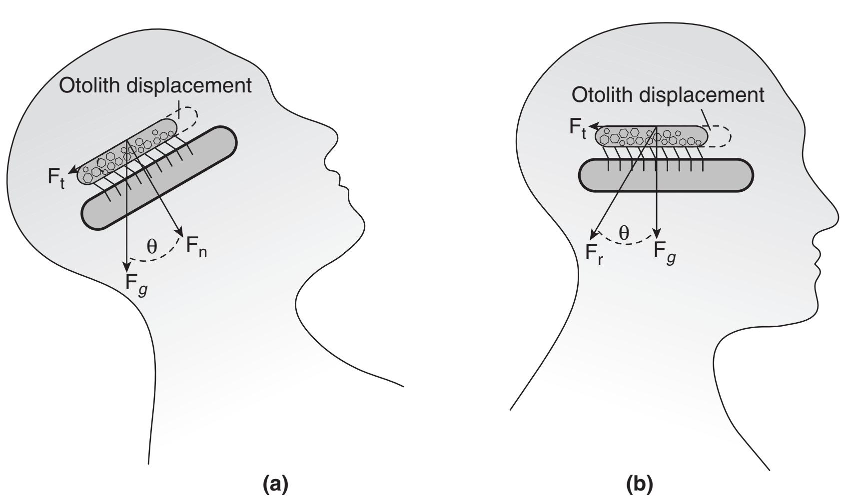


Figure 3 The otolith displacement

Notes: (a) Illustration of the otolith displacement when the head is tilted backward. F_t and F_n are the translational and normal forces generated by the force of gravity, F_g. (b) Illustration of the otolith displacement when the body is linearly accelerated forward. F_t is the inertial force, F_g is the force of gravity, and F_r is the resultant force. In both cases, the displacement angle, θ, is the same, thus signaling the same angle of head tilt.

body. Consider the following situation as illustrated in Figure 3. When the head is tilted 30° back, the force of gravity causes a backward shear on the cilia of the tilted utricular membranes, this signals a 30° head tilt to the individual. By contrast, if the head is upright and the individual is in a high performance car accelerating forward, the same backward displacement of the otolith membranes can be achieved by inertial forces, and the individual will experience backward tilt although seated upright. In this circumstance, the otolith signal is the same but the physical situations are radically different. Such illusions related to inertial accelerations are important in aviation and aerospace operations. For example, when a jet aircraft is launched from an aircraft carrier, the stimulation of the otolith organs is such that the pilot experiences a backward tilt of himself or herself and of the aircraft that is much greater than the actual angle of ascent of the aircraft. As a consequence, if the pilot is unaware of this illusion and unable to resist it, he or she can push the stick forward and drive the aircraft into the water.

The otolith organs participate in a variety of postural reflexes. Depending on the orientation of the head, the antigravity musculature of the body is differentially modulated to maintain a stable, static posture against gravity. These reflexes are thus critical for balance. Just as with the semicircular canals, the otolith organs influence the control of eye movements. For example, when the body is statically tilted in roll, it is the otolith organs that are responsible for the ocular counter-rolling, the rotation of the eyes in the opposite direction of head tilt. When the head is tilted back, the eyes deviate downward to compensate—the so-called "doll's head reflex."

Individuals who have total loss of vestibular function have to rely on vision and touch to stabilize their posture. If such individuals stand heel-to-toe and close their eyes, they will lose their balance within a few seconds. However, if they are allowed to lightly touch a stationary surface, they can stand stably even though the contact force is inadequate to provide mechanical support. The touch cues at the fingertip provide an alternative source of information about body sway, allowing balance to be stabilized. Vestibular loss individuals also have great difficulties when swimming because they can lose their orientation if they are underwater and may swim down rather than up when trying to rise to the surface. As normal individuals age, they also progressively lose receptor cells in their vestibular organs, leading to diminished postural control and increased likelihood of falling.

Motion Sickness

Virtually everyone with normal vestibular function is susceptible to motion sickness, although relative susceptibility varies enormously. Most people equate motion sickness with nausea and vomiting, but this really represents extreme sickness. Many people actually experience symptoms characteristic of motion sickness on boats, cars, and airplanes without even realizing it. For example, drowsiness, fatigue, and lack of initiative are common symptoms of low-grade motion sickness, yawning is another common sign. One observation that seems to have stood the test of time is that people without functioning otolith and semicircular canal receptors are virtually immune to motion sickness. Even if they take powerful emetic drugs such as ipecac, they do not experience nausea and vomiting, although people with normal vestibular systems will within minutes.

One of the most amazing facts about manned spaceflight is that nearly 70% of all astronauts and cosmonauts experience space motion sickness during their first three days of flight. The symptoms are brought on by head movements, especially pitch forward and backward and roll shoulder-to-shoulder movements. Yaw, side-to-side, swivel head movements are much less provocative. In spaceflight, one is weightless, which means that the otolith organs of the inner ear are unloaded. As a consequence, different orientations of the head do not result in different signals to the brain because there are no static shear forces on the otolith membrane. On Earth, pitch-and-roll head movements are accompanied by large changes in otolith activity, but yaw movements are not. This disruption of the normal relationship between head movements and otolith signals is one of the key factors in space motion sickness. After astronauts have adapted to spaceflight and are no longer experiencing symptoms, they usually experience symptoms of motion sickness upon return to Earth.

Models of Vestibular Function

It has been common since the 1950s to liken the vestibular system to inertial guidance systems that incorporate triaxial angular accelerometers and linear acclerometers. Because the otolith organs respond to gravity and inertial forces, it has been assumed that they compute the resultant vector of gravity and imposed accelerations. Under certain circumstances and depending on the frequency of linear acceleration, the central nervous system may be able to distinguish inertial from gravitational accelerations because of the brief time course of most inertial accelerations. Nevertheless, most models take the resultant vector of gravity and imposed linear acceleration as the vertical position for the organism. A new model of vestibular orientation mechanisms has been proposed recently that differs from all earlier models in that it is based on an evolutionary assumption. Organisms on Earth have evolved in the context of a static 1*g* Earth–gravity background (*g* = acceleration of Earth gravity, 9.8m/s^2). It is only recently that humans have been exposed to substantial and prolonged inertial accelerations. Before that, most imposed accelerations were associated with locomotion and voluntary movement. With the development of rocket ships, aircraft, cars, and trains, sustained exposure to non-1*g* acceleration force levels has been attained. The new model interprets shear forces on the otolith organs under normal circumstances as arising from tilt of the head in relation to gravity. In other words, the 1*g* background force level of gravity is used as a standard against which the current resultant acceleration is interpreted. The model makes important predictions for perceived orientation under conditions of unusual force levels. In particular, it predicts that for the pitch-and-roll axes, exposure to force levels greater than 1*g* will increase the amplitude of subjective tilt. But, if the individual is horizontal and tilted about the longitudinal axis (recumbent yaw), then increases in force level will not increase the apparent tilt of the body. These predictions have been validated in parabolic flight experiments involving sustained periods of 0*g* and 1.8*g* background acceleration levels. These are crucial findings because all other models of vestibular function predict that increases in force level will increase subjective body tilt for all body orientations.

James R. Lackner

See also Body Perception; Eye Movements: Behavioral; Eye Movements: Physiological; Proprioception; Self-Motion Perception

Further Readings

Bortolami, S. B., Rocca, S., DiZio, P., & Lackner, J. R. (2006). Mechanisms of human static spatial orientation. *Experimental Brain Research, 173*(3), 374–388.

Bryan, A. S., Bortolami, S. B., Ventura, J., DiZio, P., & Lackner, J. R. (2007). Influence of gravitoinertial force level during recumbent yaw axis body tilt. *Experimental Brain Research, 183*(3), 389–397.

Goldberg, M. (2000). *The vestibular system.* In E. R. Kandel, J. H. Schwartz, & T. M. Jessell (Eds.), *Principles of neuroscience* (pp. 801–815). New York: McGraw-Hill.

Lackner, J. R., & DiZio, P. (2005). Vestibular, proprioceptive, and haptic contributions to spatial orientation. *Annual Review of Psychology, 56,* 115–147.

Lackner, J. R., & DiZio, P. (2006). Space motion sickness. *Experimental Brain Research, 175*(3), 377–399.

Lackner, J. R., DiZio, P., Jeka, J. J., Horak, F., Krebs, D., & Rabin, E. (1999). Precision contact of the fingertip reduces postural sway of individuals with bilateral vestibular loss. *Experimental Brain Research, 126*(4), 459–466.

Roberts, T. D. M. (1978). *The neurophysiology of postural mechanisms* (2nd ed.). London: Butterworths.

Wilson, V. J., & Melvill Jones, G. M. (1979). *Mammalian vestibular physiology.* New York: Plenum Press.

VIBRATORY PERCEPTION

When we feel a breeze brush against our skin or when we sense the texture of a fabric while running our fingers across it, we do so, in part, on the basis of vibrations produced in our skin that are then converted into neural signals by specialized receptors. In the laboratory, the tactile perception of vibration (sometimes dubbed vibrotaction) has traditionally been investigated using sinusoidal stimuli, (i.e., stimuli that indent into and retract from the skin in regular intervals following a trajectory shaped like a sine wave). The advantage of a sinusoidal stimulus is that it can be defined using only two parameters: its frequency (the rate at which it oscillates into and out of the skin), and its amplitude (the maximum depth at which it indents the skin). The glabrous or hairless skin (of the palm and fingers, for example) has been the focus of most of the work on vibratory perception because it is with the palmar surface of the hand that we explore objects. This entry will discuss detection thresholds, the tactile perception of stimulus frequency and stimulus intensity, vibratory adaptation, and the ecological functions of vibrotaction.

Three types of receptors embedded in glabrous skin have been shown to contribute to vibratory perception: Merkel disks, which respond best to low-frequency vibrations; Pacinian corpuscles, which are exquisitely sensitive to high-frequency vibrations; and Meissner corpuscles, which are activated by vibrations at intermediate frequencies. A fourth type of mechanoreceptor, the Ruffini ending, is sensitive to skin stretch and may respond to vibratory stimuli as well, but its role in vibratory perception has yet to be elucidated.

Signals from each type of receptor are conveyed to the brain by afferent fibers, which carry information in the form of electrical impulses called action potentials. Afferent fibers are named based on the mechanoreceptor they innervate (i.e., receive input from): slowly adapting type 1 (SA1) fibers convey information from Merkel disks; rapidly adapting (RA) fibers carry information from Meissner corpuscles; and Pacinian (PC) afferents carry signals from Pacinian corpuscles. The perception of vibratory stimuli, then, depends on the production of action potentials in the three populations of mechanoreceptive afferent fibers.

Detection Thresholds

In order for a vibratory stimulus to be detected, it must evoke activity in at least one population of mechanoreceptive afferents. The sensitivity to vibration has been measured for sinusoidal stimuli spanning a wide range of frequencies. In these experiments, subjects are typically presented sequentially with two stimulus intervals, each indicated by a visual cue. A vibratory stimulus is presented in one of the two intervals, chosen at random. The subject's task is to determine whether the stimulus was presented in the first or second interval. The amplitude of the vibratory stimulus varies from trial to trial. If the subject cannot perceive the stimulus, he or she will guess and be correct 50% of the time. The objective of the experiment is to determine at what stimulus amplitude a subject will correctly identify the interval in which the stimulus was presented, at a predetermined level of performance (usually 75%); this critical stimulus amplitude is termed the absolute threshold or detection threshold. Absolute thresholds have been found to be highly dependent upon the frequency of the stimulus: At low frequencies (less than 10 hertz [Hz], i.e., less than 10 indentations and retractions per second), humans and macaques can detect stimuli as long as they are tens of microns in amplitude. As SA1 afferents are more sensitive than their RA and PC counterparts at those frequencies, signals carried by SA1 fibers mediate subjects' ability to detect these vibrations. At high frequencies (>100 Hz), absolute thresholds are much lower, sometimes as low as a tenth of a micron or less. As previously mentioned, PC afferents are exquisitely sensitive to high frequency vibrations and mediate our ability to perceive even extremely faint vibrations. At intermediate frequencies, the detection of which is mediated by RA afferents, absolute thresholds are intermediate between the two extremes (on the order of 5–10 microns). Note, however, that the sensitivity to vibrations is highly dependent upon the size of the area over which the stimulator contacts the skin, the temperature of the skin, the duration of the stimulus, and so forth. These stimulus parameters affect the afferent responses to the stimuli and thus our ability to detect them.

As previously mentioned, a sinusoidal stimulus varies along two dimensions, frequency and amplitude, each of which corresponds to a perceptual dimension. Indeed, as the frequency of a sinusoidal stimulus increases, its perceived frequency (or vibrotactile pitch) also increases, much as the pitch of an auditory stimulus increases as its frequency increases. As the amplitude of the stimulus increases, its perceived intensity (also dubbed vibrotactile loudness) increases, analogously to the concomitant increase in loudness as an auditory stimulus increases in amplitude. An important line of inquiry in somatosensory neuroscience has been to determine how mechanoreceptive afferent fibers signal the frequency and amplitude of a vibratory stimulus.

The Tactile Perception of Stimulus Frequency

Both humans and monkeys are readily capable of discriminating vibratory stimuli that differ in frequency but not in intensity. Changes of about 20% in stimulus frequency (for example, from 100 to 120 Hz) will be detected 75% of the time on average by a human observer. The ability to discriminate vibrations that differ in frequency implies that vibrations at different frequencies produce different patterns of activation in mechano-receptive afferent fibers. Indeed, afferent fibers respond in a stereotyped fashion to sinusoidal stimuli: Over wide ranges of stimulus amplitudes they produce a single action potential per indentation–retraction cycle of the stimulus. In such a response regime, they are said to be entrained to the stimulus. When the afferent is entrained to the stimulus, the time interval between successive action potentials is approximately constant as long as the stimulus is on. Thus, if the frequency of the stimulus increases, the time interval between action potentials decreases. Vernon B. Mountcastle, a pioneer in the field of systems neuroscience, and his colleagues provided strong evidence that it is the temporal patterning in the afferent response to a vibratory stimulus that signals its frequency. Indeed, at low stimulus intensities, at which afferents are not entrained to the stimulus, subjects are unable to discriminate stimulus frequency.

How information about stimulus frequency is encoded in the cortex has yet to be conclusively elucidated. Mountcastle and his colleagues initially proposed that information about frequency was conveyed in the timing of the action potentials evoked by the stimulus in neurons in the primary somatosensory cortex (SI), much as it is at the somatosensory periphery. Indeed, neurons in SI exhibit a temporally patterned response to sinusoidal stimuli, similarly to mechanoreceptive afferents. More recently, the firing rate of SI neurons has been shown to convey more information about frequency than does the precise timing of action potentials. In other words, the total number of action potentials per unit time (the total firing rate) of SI neurons changes more reliably with changes in stimulus frequency than does the time between bursts of action potentials. Determining which of these two cortical neural codes, firing rate or action potential timing, underlies our ability to distinguish stimulus frequency will require further study. Note, however, that the majority of studies investigating how stimulus frequency is represented in the cortex have been carried out using stimuli whose frequencies ranged from 10 to 50 Hz. This is an important point, as the temporal patterning in the response of cortical neurons (i.e., their entrainment) to vibratory stimuli becomes much weaker when the stimulus frequency exceeds about 80 Hz. This lack of temporal patterning stands in contrast to the ability of human observers to discriminate stimulus frequency even within this range of frequencies. How information about stimulus frequency is represented at those high frequencies remains to be elucidated.

The Tactile Perception of Stimulus Intensity

As previously mentioned, tactile stimuli do not just differ in perceived frequency, but also in perceived intensity. Again, these differences in perceived intensity imply that stimuli that differ in intensity produce different patterns of activation in the population of afferent fibers. One complication in the study of the perception of intensity is that perceived intensity is dependent upon both the amplitude and the frequency of the stimulus. An important question, then, is how information about frequency and information about intensity are disambiguated. A recent study has shown that the perceived intensity of a stimulus is determined by the firing rate it evokes in the three types of mechanoreceptive afferents. Specifically, the more action potentials are produced in the population, the more intense a stimulus will be perceived to be. Furthermore, action potentials evoked in individual fibers contribute to perceived intensity to different degrees depending on the afferent type (SA1, RA, or PC).

In SI, at least two aspects of the neural response change with increases in stimulus intensity: More neurons become activated and the most active neurons become even more active. Which of these two aspects of the cortical response determines the perceived intensity of the stimulus has yet to be determined.

Vibratory Adaptation

When a strong vibratory stimulus is presented for an extended period of time, our sensitivity to that stimulus decreases, as does the sensitivity of mechanoreceptive afferents that are activated by the stimulus. Both the perceptual and neural phenomena are dubbed vibratory adaptation. For example, after using an electric razor for a while, the skin feels numb because of this progressive desensitization of mechanoreceptive afferents located at or near the points of contact between skin and razor. In fact, adaptation not only operates on mechanoreceptive afferents but also affects neurons in the brain. Adaptation is important as it allows the sensory system to “tune out” stimuli that are constant so that we can attend to changes in our environment.

Ecological Functions of Vibrotaction

The tactile perception of vibration is important in a variety of contexts. First, the tactile detection of motion is thought to rely on the production of vibrations in the skin. Indeed, when a grasped object slips, vibrations are produced in the skin and these vibrations are then converted into neural signals, primarily by RA afferents. Second, the tactile perception of fine texture relies on vibrations. For instance, our ability to discriminate silk from cotton relies on vibratory cues: When we scan a finely textured surface, small vibrations are produced in the skin. The nature of these vibrations depends on the surface. In the absence of these vibrations (e.g., when the skin does not move relative to the surface), we are unable to distinguish its texture. Third, when we explore objects through probes, we gather information about our environment through vibrations produced in the probe. For instance, when writing with a pencil, we obtain information about the substrate upon which we write through the vibrations produced in the pencil. Similarly, when exploring with a blind cane, the vibrations produced in the cane provide information to the visually impaired individual about his or her environment. Finally, vibrations convey information about many other aspects of our environment: vibrations of a cell phone indicate an incoming call, vibrations in the steering wheel often herald engine or transmission problems, vibrations of the railroad tracks announce the impending arrival of a train, and so on.

Sliman Bensmaia

See also Audition: Loudness; Audition: Pitch Perception; Auditory Processing: Peripheral; Cutaneous Perception; Cutaneous Perception: Physiology; Haptics; Multimodal Interactions: Tactile–Auditory; Texture Perception: Tactile; Virtual Reality: Touch/Haptics

Further Readings

Bensmaia, S. J., Leung, Y. Y., Hsiao, S. S., & Johnson, K. O. (2005). Vibratory adaptation of cutaneous mechanoreceptive afferents. *Journal of Neurophysiology, 94*(5), 3023–3036.

Bolanowski, S. J., Gescheider, G. A., Verrillo, R. T., & Checkosky, C. M. (1988). Four channels mediate the mechanical aspects of touch. *Journal of the Acoustical Society of America, 84*(5), 1680–1694.

Connor, C. E., & Johnson, K. O. (1992). Neural coding of tactile texture: Comparisons of spatial and temporal mechanisms for roughness perception. *Journal of Neuroscience, 12*(9), 3414–3426.

Hollins, M., & Bensmaia, S. (2007). The coding of roughness. *Canadian Journal of Experimental Psychology, 61*(3), 184–195.

Mountcastle, V. B. (2005). *The sensory hand: Neural mechanisms in somatic sensation.* Cambridge, MA: Harvard University Press.

Mountcastle, V. B., Talbot, W. H., Sakata, H., & Hyvärinen, J. (1969). Cortical neuronal mechanisms in flutter-vibration studied in unanesthetized monkeys. Neuronal periodicity and frequency discrimination. *Journal of Neurophysiology, 32*(3), 452–484.

Muniak, M. A., Ray, S., Hsiao, S. S., Dammann, J. F., & Bensmaia, S. J. (2007). The neural coding of stimulus intensity: Linking the population response of mechanoreceptive afferents with psychophysical behavior. *Journal of Neuroscience, 27*(43), 11687–11699.

Salinas, E., Hernandez, A., Zainos, A., & Romo, R. (2000). Periodicity and firing rate as candidate neural codes for the frequency of vibrotactile stimuli. *Journal of Neuroscience, 20*(14), 5503–5515.

Talbot, W. H., Darian-Smith, I., Kornhuber, H. H., & Mountcastle, V. B. (1968). The sense of flutter-vibration: Comparison of the human capacity with response patterns of mechanoreceptive afferents from the monkey hand. *Journal of Neurophysiology, 31*(2), 301–334.

VIDEO GAMES

Over the past 20 years, the percentage of Americans who play *video games* has grown at an astronomical pace; 59% of all Americans and 72% of Americans ages 6 to 44 currently play video games. A growing body of research indicates that individuals who play a particular subset of video games, dubbed action video games, demonstrate marked enhancements in a number of sensory/cognitive skills. This entry describes action video games, the effect of action video games on perceptual skills, demonstrating causation between video game experience and sensory/cognitive skills, and real-life applications of video game experience.

Action Video Games

For purposes of this entry, action video games are defined as those games that have many quickly moving objects, that require effective peripheral processing as items in the periphery must constantly be localized and identified, and where the number of independent items that need to be kept track of far exceeds the circumstances experienced in normal life. In short, an action video game is one that places extraordinary demands on the visual and visuo-motor systems. Games from a variety of genres are included as action games, but most typically they include first-person shooters (such as Halo), third-person shooters (such as Gears of War), and some car racing games (such as Burnout). It is crucial to note that not all types of video games can be expected to affect processing in the same manner or to the same degree. For instance, as fantasy games are typically much less visually demanding than action games, they are unlikely to generate the same types of perceptual changes.

Effect of Action Video Games on Perceptual Skills

Action video game experience has been shown to improve a number of reasonably fundamental aspects of visual perception. Different aspects of vision have been shown to be enhanced as a result of action video game play, including spatial visual search (finding a target among distracters, as in finding a set of keys on a cluttered desk), temporal visual search (finding a target when it is presented in a stream of images one after another, as in seeing a particular actor when you rapidly flip through television stations), the number of independently moving items that can be concurrently tracked (as in keeping track of five children running on the playground who are surrounded by identically dressed "distracting" children), dividing attention between two locations, mentally rotating objects, dual-task performance, and the spatial resolution of vision (the ability to resolve extremely small items even when other distracting items are crowded around). There are also myriad cases wherein action video game experience leads to substantially faster response times to visually presented information. Although the majority of the work has been done using young adult (college-age) individuals, similar effects have been noted in both young children and the elderly.

Demonstrating Causation

Typically in these experiments, performance is first contrasted between avid action video game players (VGPs—individuals who play 5 or more hours a week of action video games) and nonaction video game players (NVGPs). If, as in the experiments noted earlier (visual search, tracking objects, etc.), VGPs are seen to outperform NVGPs, a second step must be taken to demonstrate that a causative relationship exists. After all, it could simply be the case that individuals who are born with better perceptual skills are more successful at video games and thus tend to play them more than individuals who are born with poorer perceptual skills. To address this question, groups of NVGPs are specifically trained on action video games (anywhere from 10 to 50 hours). Their performance on the measure of interest (visual search ability, for instance) is assessed a few days before and after training. These results are then compared to a control group that is trained for the same duration on a nonaction video game, which allows the researcher to control for test–retest improvements (one would expect someone to improve on a test the second time he or she takes it regardless of intervention) and for psychological effects, such as the Hawthorne effect (subjects who are watched and who have an interest taken in their performance tend to improve their performance). In the cases previously listed, those individuals trained on action games were seen to improve by greater margins than the control group, thus establishing that action video game experience causes enhancements in perceptual ability.

Real-Life Applications of Video Game Experience

Professional

As early as the 1980s, the U.S. military had begun to explore the potential uses of video games as a training aid, finding that video game skill was a powerful predictor of aircraft carrier landing abilities and that video games were an excellent candidate for a performance test battery. In the 1990s, Gopher and colleagues demonstrated that training cadets on a video game led to large improvements in flight scores. More recently, the U.S. Marine Corps and U.S. Army have developed their own video games to recruit and train soldiers.

In addition to military applications, another group that has shown benefits from video game training is laparoscopic surgeons. Several recent reports have suggested that extensive action video game experience is correlated with faster surgical completion times and fewer surgical errors.

Rehabilitative

In many cases, the elderly suffer from natural declines in the same types of abilities that are enhanced by action video game experience in young adults. Not surprisingly therefore, research has indicated that action video game play in the elderly leads to increases in speed of processing, manual dexterity, hand-eye coordination, response selection, and general cognitive abilities, such as short-term memory and reasoning.

Amblyopia is a developmental visual disorder that is a result of problems with neural processing (as compared to standard nearsightedness or farsightedness that is a result of imperfections in the shape of the eyeball). While typical treatments for amblyopia may be considered dull and thus can suffer from poor compliance, video games may offer a more exciting and thus more successful rehabilitation program.

Daphne Bavelier and C. Shawn Green

See also Amblyopia; Attention: Covert; Attention: Effect on Perception; Attention: Selective; Attention: Spatial; Computer Graphics and Perception; Experience-Dependent Plasticity; Perceptual Learning

Further Readings

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*(10), 850–855.

Gopher, B., Weil, M., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. *Human Factors, 36,* 387–405.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534–537.

Green, C. S., & Bavelier, D. (2006). The cognitive neuroscience of video games. In L. Humphreys & P. Messaris (Eds.), *Digital media: Transformations in human communication* (pp. 211–224). New York: Peter Lang.

Trick, L. M., Jaspers-Fayer, F., & Sethi, N. (2005). Multiple-object tracking in children: The "Catch the Spies" task. *Cognitive Development, 20*(3), 373–387.

VIRTUAL REALITY: AUDITORY

Virtual reality (VR) systems provide users with a computer-mediated experience of a sensory world that exists for them as a virtual world within which they can act and interact. Popular fiction has portrayed the VR experience as generating such a realistic illusion that users can forget that they are experiencing a virtual world that is generated by computer. In practice, however, such strong illusions do not typically result unless users are presented with a coordinated interactive display of information via multiple sensory modalities. Therefore, VR systems often include auditory displays as well as visual displays so that users can hear as well as see what occurs in the virtual environments that are generated for them through computer simulations. Of course, haptic display technology can also be an extremely important component of a VR system because it supports the user experience of touching and feeling virtual objects; but here again, the illusion of reality will be best supported if an appropriate sound is heard, for example, whenever the user's hand makes contact with a virtual object. This entry covers spatial auditory displays, perception and action in auditory virtual reality, and research in auditory virtual reality.

Consider the potential advantage users might enjoy if they are able to receive coordinated multimodal stimulation in a *telerobotic* application, such as remote operation of a vehicle through unknown terrain. Driving a vehicle remotely will almost certainly be facilitated if users are able to listen to sound arriving from all directions around them while focusing their gaze on the environment to be navigated via a relatively constrained visual field. For example, imagine how difficult it would be to drive a car on ice and drifting snow without being able to hear or feel the change in the traction of the tires. Such auditory and haptic feedback could well guide user behavior, allowing successful navigation when vision alone is not enough.

Of course, there are many VR systems serving applications that do not include auditory displays, such as VR-based scientific visualization systems. These are systems that operate upon sets of data that exist in three or more dimensions and can represent the datasets as solid objects that users can explore visually, without any sonic component. On the other hand, VR-based entertainment systems as a rule will include audio, and many use sophisticated audio signal processing to create realistic spatial impressions of virtual sound sources positioned in the virtual environments in which users find themselves. For example, immersive VR-based games use spatial auditory display technology to provide users with an enhanced awareness of their situation. Indeed, just as in everyday life, auditory events occurring outside of the field of view often lead users to direct their gaze to bring the source of the auditory event into focal vision.

Spatial Auditory Displays

Spatial auditory display technology is used to produce, at the ears of VR-system users, sound signals that closely match the signals that would reach their ears if the sound sources were present in an actual acoustic environment similar to that which the virtual environment simulates. The common approach to such virtual acoustic simulation is to use headphone reproduction of sound signals created through binaural synthesis, although multichannel loudspeaker systems also can create convincing illusions of spatially immersive virtual acoustic environments. Binaural synthesis uses measured or simulated *head-related transfer functions* (HRTFs) that capture how impinging sounds are transformed by the acoustics of a listener's head and external ears. These HRTFs impose acoustic cues on sound sources that enable listeners to localize those sources in three-dimensional (3-D) space. Although source movement to the left and right of the listener can be cued by changing the relative level of the sound at the two ears, HRTFs can enable distinctions between sources moving forward and backward or upward and downward in space. These HRTFs can be measured for each individual by placing small microphones inside their ears, and the results of these measurements stored for subsequent use in binaural sound processing for that individual. Such binaural synthesis has been commercially available via real-time digital signal processing since the early 1980s, but there remains some controversy surrounding the issue of whether or not the best HRTF-based spatial auditory display results require the use of the HRTFs measured for the individual user. Generic transfer functions that offer a generalized solution to the problem of sound *spatialization* may be adequate for most users in many VR applications, and contemporary binaural synthesis solutions often provide a means for customizing the transfer functions according to the anatomical size of the user.

Perception and Action in Auditory Virtual Reality

In most applications of auditory VR technology, interactive spatial sound processing is required for satisfying results; that is, the computer-generated sound signals should create illusions of virtual sound sources that move relative to the user, either when the sources or the user (or both) change their location or orientation over time. If sound is reproduced using headphones as part of a *head-mounted display* (HMD) system, it is most common to use HRTFs in an interactive spatial auditory display. Such a display uses head-tracking technology to update the simulation parameters during user motion so that stationary virtual sound sources are perceived to remain in spatially stable positions within the virtual environment. This becomes extremely important in resolving potential confusion between sound sources located to the front or the rear of the listener, because rotation of the user's head has opposite effects on the arrival time of the displayed signals for these two directions: A frontal source is brought closer to the right ear by a left turn of the user's head, whereas a source located to the rear is brought closer to the left ear by that same left turn of the user's head.

Of course, if a spatial auditory display system uses many loudspeakers distributed around the user in three dimensions (i.e., an extension of consumer surround sound systems), then virtual sources that are to be localized behind the user may be reproduced by loudspeakers that are also located behind the user, and so tracking changes in the user's head position and orientation becomes less critical. Although such "many-channel" loudspeaker arrays have some definite advantages over head-mounted auditory displays, only a small number of such expensive multichannel systems are in use, typically as relatively permanent installations in dedicated research labs (which can feature hundreds of small loudspeakers configured within a spherical array).

More sophisticated spatial auditory displays attempt to faithfully simulate the reverberation of sound within the enclosed space in which the virtual sound sources are located. This requires the simulation of indirect sound that is reflected by walls and other surfaces in the space. Indeed, for realistic results, the direction in which a sound source is facing relative to the observer should also have an impact upon the simulation of sound propagating through the virtual environment. In some demanding architectural applications, attempts have been made to base the sound simulation on highly detailed models with varying sound absorption characteristics for each reflecting surface in the virtual acoustic space. Real-time interactive rendering in such cases requires considerable processing power, because the indirect sound must be perceived as arriving from different directions.

Research in Auditory Virtual Reality

In contrast to the primary motivation for most spatial hearing research, which has been to gain greater understanding of the mechanisms of human spatial hearing, the motivation for most research in auditory VR has been to establish the adequacy of various spatial auditory display technologies intended to synthesize auditory scenes for human users. In VR-based computer games, for example, detailed architectural information is often ignored, and greater amounts of computational power may be focused upon displaying more simultaneous virtual sources along with their local sonic interactions, such as sound diffracted around nearby obstructions and reflected from nearby surfaces. As the number of virtual sound sources grows in such virtual worlds, resource allocation becomes a concern, and perceptual factors such as masking can be considered in order to minimize computational cost while maximizing effectiveness of the spatial auditory display

William L. Martens

See also Acoustics and Concert Halls; Auditory Localization: Psychophysics; Multimodal Interactions: Tactile–Auditory; Multimodal Interactions: Visual–Auditory; Sound Reproduction and Perception; Virtual Reality: Touch/Haptics; Virtual Reality: Vision

Further Readings

Begault, D. R. (1994). *3-D sound for virtual reality and multimedia.* Boston: Academic Press Professional.

Blauert, J. (2005). Analysis and synthesis of auditory scenes. In J. Blauert (Ed.), *Communication acoustics* (pp. 1–25). Berlin: Springer-Verlag.

Martens, W. L. (2003). Perceptual evaluation of filters controlling source direction: Customized and generalized HRTFs for binaural synthesis. *Acoustical Science and Technology, 24*(5), 220–232.

Novo, P. (2005). Auditory virtual environments. In J. Blauert (Ed.), *Communication acoustics* (pp. 277–297). Berlin: Springer-Verlag.

VIRTUAL REALITY: TOUCH/HAPTICS

Haptic technology does for the sense of touch what computer graphics do for vision. Haptic technology creation of computer-generated haptic virtual objects (HVOs), which can be touched and manipulated with one's hands or body. HVOs provide a rich combination of cutaneous and kinesthetic stimulation through a bidirectional haptic (touch) information flow between HVOs and human users.

Many mechanical properties of everyday objects are experienced through touch. These properties include weight and shape of objects, object elasticity, object's surface texture (e.g., smooth or rough), and so forth. HVOs can have many of these real-object mechanical properties. Perhaps more importantly, HVOs can have mechanical properties that do not exist in nature. For example, HVOs can possess paradoxical, normally impossible, combinations of mechanical properties. People touching such paradoxical HVOs can experience surprising perceptual effects. For example, people actually touching a surface with a hole can instead haptically perceive a surface with a bump on it. Human haptic capabilities can be investigated in totally new ways with these and other HVOs. HVOs are also used in touch-enabled human-machine interface applications. For example, HVOs are used to create virtual internal organs that can be touched and manipulated by surgical trainees. This entry discusses the creation of haptic virtual objects and using them as experimental stimuli.

Creation of Haptic Virtual Objects

Generally, HVOs are created through force fields (or "force-feedback"), generated by computer-controlled mechanical systems called haptic interfaces (HIs). An HI delivers the force-feedback to a person's hands or body. This reproduces major aspects of what actually happens when touching real, everyday objects.

For example, in Figure 1(a), a person handles a stick to poke and deform a real (not an HVO), flexible surface (e.g., a rubber sheet). Following the physics of this mechanical interaction, the surface exerts a force back into the stick. This contact force is transmitted to the person's hand. The person experiences this force as the surface's resistance to deformation. This interaction scenario can be reproduced with an HVO. For this, the person holds an HI's sticklike manipulandum (Figure 1b). The rest of the HI mechanism is only partially shown (the bar connected to the manipulandum's tip, Figure 1b). The person moves the manipulandum in an empty, delimited, three-dimensional space (the HI workspace). HVOs are created within this workspace. The person uses the manipulandum to "poke" an HVO (dashed surface, Figure 1b). The HVO is not a physical object at all. It consists only of computer-controlled forces that are generated as follows.

As the person moves the manipulandum, HI's sensors measure the current position of the manipulandum's tip (Figure 1b). A control computer (CC) monitors this position. When the person moves the manipulandum's tip into the workspace region occupied by the HVO, the CC detects this "collision" of the manipulandum's tip with the HVO. Then, the CC calculates a simulated contact force from a model (e.g., equations) of the real interaction's physics. Next, the CC activates the HI's actuators (e.g., electric motors) that, in combination with HI mechanics, produce an actual physical force that is applied into the manipulandum's tip (Figure 1b). This force physically realizes the simulated contact force. As when touching the real surface (Figure 1a), the person feels HI forces through the manipulandum. As the person explores and even "deforms" the HVO, adequate HI forces sculpt, so to speak, the HVO. This software-controlled HVO creation process is called haptic rendering (HR). The CC executes HR events (collision detection and force calculation and generation when necessary) at a high rate (1 kHz or more). In contrast, graphics rendering typically works at a rate of 30 Hz.

Different manipulandums can be used to interact with HVOs, such as tools resembling thimbles (for fingertip insertion), scissors (for surgical simulation), and hand exoskeletons. Multifinger HVO interaction is also possible. HVOs can be combined with visual, auditory, and other display technologies for multimodal sensory stimulation.

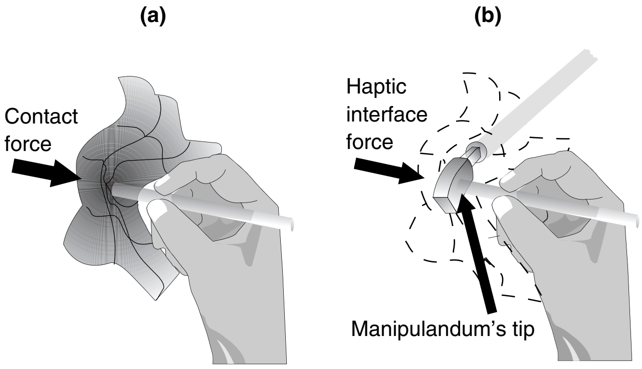


Figure 1

Notes: (a) Poking a real flexible surface with a stick (b) Holding an HI manipulandum.

Haptic Virtual Objects as Experimental Stimuli

HVOs' mechanical properties can be selectively defined and changed through software. This is difficult (or impossible) to achieve with real objects. Therefore, HVOs allow entirely new ways to investigate perception. For example, people touching paradoxical HVOs can experience contact forces normally found when touching a surface with a bump, while actually touching a surface with a hole. As a result, people haptically perceive a surface with a *bump*. Such situations reveal that contact forces can determine how object shape is haptically perceived, even when other sensory cues (e.g., the object's actual geometry) offer conflicting haptic information. The full potential of HVOs remains to be explored in fields such as haptic perception, multimodal and intermodal perception and integration, motor control and dexterous manipulation, functional brain imaging (e.g., for correlation of brain physiology with real-time HVO interaction), and in applications such as surgical simulators and touch-enabled Internet browsing.

For effective use of HVOs in experiments and more generally, it is essential to thoroughly understand (a) the physics of the haptic interaction scenario of interest; (b) the capabilities and limitations of available haptic technology; (c) how the haptic interaction scenario and its physics can be implemented with haptic technology; (d) how this implementation may differ from the original scenario, and how this may influence human perception and performance; and (e) how judicious experimental design and testing can deal with all these issues.

Gabriel Robles-De-La-Torre

See also Cutaneous Perception; Haptics; Kinesthesia; Multimodal Interactions: Visual–Haptic; Texture Perception: Tactile; Vibratory Perception

Further Readings

Basdogan, C., & Srinivasan, M. A. (2002). Haptic rendering in virtual environments. In K. Stanney (Ed.), *Handbook of virtual environments* (pp. 117–134). London: Lawrence Erlbaum.

Hayward, V., Astley, O. R., Cruz-Hernandez, M., Grant, D., & Robles-De-La-Torre, G. (2004). Haptic interfaces and devices. *Sensor Review, 24*(1), 16–29.

Robles-De-La-Torre, G. (2006). The importance of the sense of touch in virtual and real environments. *IEEE Multimedia, 13*(3), 24–30.

Robles-De-La-Torre, G. (2008). Principles of haptic perception in virtual environments. In M. Grunwald (Ed.), *Human haptic perception: Basics and applications*. Basel, Switzerland: Birkhäuser Verlag.

Robles-De-La-Torre, G., & Hayward V. (2001). Force can overcome object geometry in the perception of shape through active touch. *Nature, 412*(6845), 445–448.

VIRTUAL REALITY: VISION

Virtual reality systems (often referred to simply as VR) are computer interfaces that provide users with the sensory experience of being in a simulated space. The goal is to provide a person with a compelling sense of being present in a computer-generated *virtual world* that is too expensive, too dangerous, or not practical to visit in person. A multiplicity of VR-based entertainment systems currently exist, ranging from video games to amusement park rides. VR is also playing an increasing role in applications such as training, mission planning, education, and rehabilitation, where it is important that users act in a virtual world in the same manner as they would if placed in the corresponding physical world that is being simulated. This entry describes visual displays, perception and action in virtual reality systems, and the use of virtual reality for studying perception.

Visual Displays

Ideally, a virtual reality system would provide synthesized stimuli for all of the senses, and there is debate about whether or not effective VR is possible without providing a full range of sensory modalities. In practice, however, the majority of existing VR systems are limited to visual displays, using computer graphics to create views of the synthetic world. Two properties of these VR visual displays seem essential to approximating a perceptual sense of being in a simulated world. *Visual immersion* excludes visual sensation of anything except the simulation itself. This avoids

Figure 1 Head-Mounted Display

sensory and perceptual conflicts between the real and virtual worlds, increasing the impact of even low fidelity virtual displays. *Head tracking* involves generating the computer graphics presented to the user from a viewpoint that changes appropriately as the user moves. This provides motion parallax cues allowing perception of the spatial structure of the virtual environment based on the changing view a result of changing viewpoint, but more importantly provides a compelling sense of being "in" the simulation.

Two classes of visual displays are used in virtual reality systems, each with distinct advantages and limitations. *Head-mounted displays* (HMDs) are devices worn on the head and feature two small display screens, one in front of each eye (Figure 1). Computer graphics are used to generate the video fed to each of these screens. Lenses (*collimated optics*) are used to cause the screens to appear farther away than they really are. This allows a user's eyes to focus (accommodate) to a distance consistent with a large space and makes the viewing of displayed images less sensitive to the precise positioning of the HMD on the user. The field of view of HMDs is usually quite limited, ranging horizontally from about 20° to about 45°, though a few HMD systems have fields of view approaching normal human vision. When used in virtual reality simulations, head-mounted displays incorporate a tracking device that measures the position and orientation of the HMD and feeds this information to the computer graphics system. The computer graphics system generates imagery that would be seen in the virtual world from the corresponding position and direction of view.

In *screen-based systems,* views of the simulated world are displayed on large, fixed, back-projected screens rather than on minidisplays worn by the user. The most effective screen-based VR systems surround the user with screens, sometimes including the ceiling and floor. VR systems involving a single screen are also used, though it is difficult with such systems to achieve a high level of visual immersion. Head tracking, usually involving the user wearing some sort of simple target on her or his head, allows the computer graphics projected onto the display screens to change in response to the user's movements.

Stereo image display is usually used with head-mounted displays and is common in many screen-based displays. In either case, the computer graphics system needs to generate two different video streams from two slightly offset positions, one for each eye. In the case of head-mounted displays, these two video streams are simply fed to the two different display screens worn by the user. One way for creating screen-based stereo is to alternately project left-eye and right-eye images on the screen and then use synchronized shutter goggles to alternate viewing of the screen by the left and right eyes. Another common way mixes left eye and right eye images into a single image but with different polarizations, with the viewer wearing appropriately polarized goggles.

Perception and Action in Virtual Reality Systems

Many HMD and screen-based VR systems allow a user to naturally walk while viewing the virtual world. In the case of HMDs, walking is limited to the extent of the space over which head tracking can be done, which typically ranges from a few meters to a few 10s of meters. Screen-based systems allow walking over a maximum of a few meters. When natural walking is supported in a virtual reality system, visual information for self motion, which comes from the computer graphics being coupled with head tracking, is augmented with proprioceptive and vestibular information arising from real movement. Natural walking through a virtual environment with the visual display changing in a manner appropriate to actual movement generates a compelling sense of moving through the virtual world.

Some virtual reality systems include hand-held controllers that are tracked by the same system used for head tracking. These controllers, often called wands, can be used to point at locations in the virtual world. An image of the wand, in the appropriate location relative to the user, is sometimes rendered in the computer graphics. This seems to help in positioning the wand and in adding a better sense of realism in the wand usage. A similar idea is incorporated into the Nintendo Wii video game system, in which the position of a hand-held remote control is tracked in such a way that it can function as many different virtual objects, ranging from a tennis racket to a boxing glove.

In principle, a properly designed virtual reality system should provide the same visual cues for space perception as the real world, with the exception of accommodation, which is fixed in such systems, and the more limited field of view. In practice, fixed accommodation leads to difficulties in binocular convergence (a depth cue based on comparing the line of sight from the two eyes to a location of interest), and so can distort distance judgments based on stereoscopic vision, particularly at near distances. Research on distance perception in VR over reaching distances out to 1 to 2 m from the viewer indicates that distance judgments are distorted from what is intended in the model of the virtual world. The nature of this distortion is idiosyncratic to the specifics of the display, however, perhaps because stereo plays such an important role in depth perception over this range of distances. Between 2 and 20 m, distance judgments in HMD-based virtual reality displays show consistent underestimation when evaluated using the same measures that show near-correct performance in the real world, with accuracy reported as 50 to 85% of the intended distances. There has been much speculation as to why this is so, but the cause has not yet been identified. Space perception has been studied less in screen-based systems, though anecdotal evidence suggests a similar underestimation of depth. There has been little study of space perception in virtual reality involving judgments other than of depth.

Use of Virtual Reality for Studying Perception

Virtual reality systems are increasingly used as a tool in basic perceptual research. Perception experiments done using traditional methods often require simplified or unrealistic stimuli in order to maintain sufficient experimental controls. This can limit the ecological validity of the results obtained. VR has the potential to generate experimental conditions that are simultaneously realistic and well controlled. In addition, the very nature of VR allows experiments to be conducted using stimuli that are not physically realizable in the real world. For example, visual information can be presented that is inconsistent with a person's own movement or with what a person feels through touch, allowing investigations of the integration of different sensory modalities by systematically disassociating different cues in a VR display.

William B. Thompson

See also Action and Vision; Depth Perception in Pictures/Film; Navigation Through Spatial Layout; Pictorial Depiction and Perception; Spatial Layout Perception, Psychophysical; Virtual Reality: Auditory; Virtual Reality: Touch/Haptics; Visual Scene Perception

Further Readings

Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers, 31*(4), 557–564.

Loomis, J. M., & Knapp, J. (2003). Visual perception of egocentric distance in real and virtual environments. In L. J. Hettinger & M. W. Haas (Eds.), *Virtual and adaptive environments* (pp. 21–46). Mahwah, NJ: Erlbaum.

Sutherland, I. E. (1968). A head-mounted three dimensional display. In *Proceedings of the fall joint computer conference* (pp. 757–764). Washington, DC: Thompson.

Whitton, M.C. (2003). Making virtual environments compelling. *Communications of the ACM, 46*(7), 40–47.

VISCERAL PERCEPTION

Visceral perception refers to the ways people feel sensations in their bodies. Perceptions of a stomach ache, a racing heart, or the feelings of needing to urinate are common examples of visceral perception. Historically, it was thought that our internal organs reliably sent signals to our brain so that we could accurately detect the states of our bodies. In recent years, however, it has become clear that our ability to read and understand our bodily cues is as complicated as our abilities to perceive sights and sounds from the outside world.

Visceral perception is only one kind of bodily perception, or *interoception.* While sitting in your chair reading this article, your brain is processing all sorts of bodily signals. Your ability to "know" that you are sitting upright and maintaining your balance is called *proprioception.* Proprioception requires that signals from your muscles, joints, and even the balance centers from your inner ears are sent to your brain. Both proprioception and visceral perception can occur with or without conscious awareness, just as with perception of events that occur outside the body.

Visceral perception lies at the heart of many psychological and medical issues. For example, our ability to know if we are angry or sad depends, in part, on how our bodies feel. Knowing if you are sick, hungry, or sexually aroused depends on the accurate reading of our internal organs. Indeed, many physicians have become frustrated when patients complain about feelings of irregular heartbeats, shortness of breath, or other symptoms when their hearts, lungs, or other organs have been functioning properly. The problem, then, is that the patients may have distortions in their abilities to accurately perceive bodily activity. That is, their visceral perception abilities have failed in some way.

This entry discusses the sensory bases for visceral perception, limits of the conscious awareness of visceral activity and methods for measuring accuracy in visceral perception, and the value of visceral perception for understanding health problems such as noncardiac chest pain and irritable bowel syndrome.

Receptor Mechanisms of Visceral Sensation

The internal organs have elaborate nerve cells that can sense movement and pressure (*mechanoreceptors* and *baroreceptors*), chemicals (*chemoreceptors*), temperature (*thermoreceptors*), and painful stimuli (*nociceptors*). These nerve receptors associated with vision, hearing, taste, and smell are assembled into localized structures such as the eyes and ears. In contrast, receptors for sensing visceral events are distributed throughout the body. These receptors may be categorized into two types: (1) rapidly adapting in response to changes that occur in the organ, and (2) slowly adapting receptors that are sensitive to the organ's ongoing state.

Among the most-studied visceral receptors are the *baroreceptors,* which sense changes in blood pressure. One type of baroreceptor is found in the aorta and carotid arteries and is sensitive to increased blood flow; when blood flow increases, these arterial receptors send signals to the brain to reduce blood flow from the heart so that blood pressure does not escalate out of control. Some researchers believe these rapidly adapting receptors are primarily responsible for subjective feelings of heart activity. Another baroreceptor type, one that reflects slow adaptation, is in the walls of the heart and the veins. It alerts the brain when the volume of blood drops below a desired level, which in turn can trigger circulating hormones to rebalance the water and salt content of the blood and thereby raise overall blood pressure.

Receptors are located throughout the stomach and gastrointestinal tract. Mechanoreceptors sensitive to swelling of the stomach are generally believed to be related to feelings of fullness and to food intake. When the stomach is empty, people who have swallowed a gastric balloon that is then fully inflated eat less once the balloon is deflated, compared to when the balloon is inflated to lesser volumes, suggesting a threshold for the effects of distension. In addition, people's ratings of hunger decrease and ratings of fullness increase as the balloon is inflated. On the other hand, if the gastric balloon is not just deflated but *removed* before the opportunity to eat, the amount of premeal distension has little effect on the amount eaten. These results imply that the mechanoreceptors of the stomach are of the rapidly adapting type.

Visceral receptors aren't isolated from other receptors located inside the body. For example, although baroreceptors in the aorta contribute to awareness of heartbeat sensations, receptors in the muscles of the chest can also be stimulated by heart actions. The stronger the heartbeat, the more likely the heart may make contact with chest muscles or the sternum. Likewise, gastric distension or intestinal movements can bring the internal organs into contact with abdominal muscles. In both cases, movement signals may be picked up by other receptors in the body. As we pay attention to our bodies, any and all of these bodily signals will be detected by our brains, which will lead to a general "feeling" that may or may not accurately reflect what is going on in a particular location in our body.

In fact, things like body posture can influence people's ability to detect visceral events, such as heartbeats. Heartbeat discrimination improves as the body moves from standing to lying down; it improves even more when the body is prone and the heart makes direct contact with the chest. Body fat can make it more difficult to feel many visceral events, including heartbeats and gastric contractions. For example, men are often better at accurately detecting their heartbeats and stomach contractions than women. This may be partially due to the fact that women have more fat in comparison to muscle than do men.

Conscious Awareness and Detection Accuracy

Conscious awareness of bodily states can guide behavior, as when sensing a full bladder makes us search for a restroom, the feeling of a full stomach makes us stop eating, and so forth. Most of the time, our internal organs do not occupy attention; when attention *is* attracted, it can serve as a cue to action. However, visceral perception need not be conscious; most is not. Visceral behaviors can become classically conditioned through association with an external stimulus; such conditioning may occur without awareness of the visceral reaction. For example, the sight of a feared object, such as a person we consider a bully (a conditioned stimulus), may trigger changes in cardiovascular reactions, which do not attain consciousness; instead, our conscious experience is the desire to avoid the bully, and subsequent behaviors may be directed toward fulfilling this conscious desire. In everyday language, we probably describe the experience using emotion words, not in bodily terms (although there is inevitably much overlap between the two classes of usage). In this case, conditioned cardiovascular responses contribute to our conscious experience but do not constitute it.

Although it is not necessary to be aware of one's visceral reactions for these to have meaningful effects on experience and behavior, it is nonetheless of great interest to know how accurately people can detect changes that occur inside the body. We use bodily imagery to describe our emotional experiences ("My heart jumped into my throat!") and illnesses ("I feel this pressure in my stomach"). To what extent are we referring literally to measurable visceral responses that we have sensed? There is no simple answer to this question. Visceral organ systems differ in the number and location of sensory receptors, and measures of organ activity and function do not necessarily have ordinary-language synonyms. To map the correspondence between bodily responses and people's reports of those responses, investigators of visceral perception accuracy have created a variety of methods to measure the ability to detect discrete physiological events, such as heartbeats, and changes in physiological state, such as distension of the stomach. Such research is useful for understanding both accuracy and the larger matter of consciousness in visceral perception.

Ultimately, people vary widely in their abilities to detect visceral changes in their bodies. Whether the task is to press a button in approximate rhythm with heartbeats, distinguish which of two stimuli was coordinated with heartbeats, state whether a gastric or rectal balloon has been inflated or deflated, or rate breathing symptoms in response to inhalants, individual performance lies along a continuum ranging from very inaccurate to quite impressive. Similarly, people's performance on detection tasks depends on the task itself; the effect of bodily position on cardiovascular perception is an example. More importantly, people who accurately perform one detection task are neither more nor less likely to be accurate when tested using a different technique, even when both tasks ostensibly measure sensitivity to the same physiological events.

Visceral perception studies point to at least two broad questions. First, *what* do people perceive when they are asked to report on their sensations? That is, when asked to count their heartbeats, for example, are individuals sensing heartbeats per se or some other aspect of cardiovascular activity, such as the strength of ventricular contractions and the consequent volume of blood ejected from the heart? Current evidence indicates it is the latter.

A second important question raised by studies of visceral detection is, are such measures meaningful for addressing clinical problems? Are abilities to detect isolated physiological reactions in a controlled laboratory setting relevant for understanding (say) illness complaints?

Visceral Sensitivity and Illness Complaints

Chest pain, stomach pain, and intestinal dysfunctions are common illness complaints that are sometimes difficult to diagnose based on physical signs. Noncardiac chest pain (NCCP) and vague stomach pain, such as functional dyspepsia (FD) and irritable bowel syndrome (IBS), are *diagnoses of exclusion,* made after the search for organic damage has yielded no clear findings. The sensitivity of individuals with these diagnoses to controlled stimulation of the gastrointestinal tract has been extensively tested. People with NCCP are more sensitive to balloon distension and mild electrical stimulation in the esophagus, sometimes reporting changes in such stimuli when none has occurred—*false alarms.* The tendency to react to lower levels of balloon distension in the stomach has been observed in people with FD, and a similar sensitivity to rectal and colon distension occurs in IBS; there also is greater likelihood of false alarms. Such sensitivity might result from differences in receptor activity, in the afferent nerves that conduct such activity to the central nervous system, in aspects of central nervous system processing of afferent signals, or in all of these. False alarm observations imply that some people with these disorders have a bias when processing visceral information.

Similar biases have been recently reported in individuals with very different health problems. Some people newly diagnosed with asthma report feeling more airway obstruction and fatigue after inhaling a saline air mixture than room air, even though the saline causes no changes in measures of pulmonary function. Sensitivity to saline was most likely among those who scored high on the trait of negative affectivity (NA), the tendency to report negative emotions. Among people with congenital heart disease, self-reports of cardiac symptoms during a stressful task increase only for individuals who are highly anxious, not those who are low in dispositional anxiety. It is fair to call the increase in cardiac symptom reporting "biased" because the task actually causes no change in measured cardiovascular responses.

Because these tendencies are evident across illnesses in several organ systems, the further implication is that they represent a durable response bias in visceral perception, not a peculiarity of certain individuals or illnesses. In summary, measures of visceral detection accuracy can add to understanding an individual's perceptions of health problems.

James A. Skelton and James W. Pennebaker

See also Body Perception; Cutaneous Perception: Kinesthesia; Mind and Body; Pain: Assessment and Measurement; Proprioception; Unconscious Processes

Further Readings

Ádám, G. (1998). *Visceral perception: Understanding internal cognition.* New York: Plenum Press.

Cameron, O. G. (2002). *Visceral sensory neuroscience: Interoception.* New York: Oxford University Press.

De Peuter, S., Put, C., Lemaigre, V., Demedts, M., Verleden, G., & Van den Bergh, O. (2007). Context-evoked overperception in asthma. *Psychology & Health, 22*(6), 737–748.

Karsdorp, P. A., Kindt, M., Rietveld, S., Everaerd, W., & Mulder, B. J. M. (2007). Stress-induced heart symptoms and perceptual biases in patients with congenital heart disease. *International Journal of Cardiology, 114*(3), 352–357.

Mertz, H., Fullerton, S., Naliboff, B., & Mayer, E. A. (1998). Symptoms and visceral perception in severe functional and organic dyspepsia. *Gut, 42*(6), 814–822.

Pennebaker, J. W. (1982). *The psychology of physical symptoms.* New York: Springer-Verlag.

Skelton, J. A., & Croyle, R. T. (Eds.). (1991). *Mental representations in health and illness.* New York: Springer-Verlag.

Whitehead, W. E., & Palsson, O. S. (1998). Is rectal pain sensitivity a biological marker for irritable bowel syndrome: Psychological influences on pain perception. *Gastroenterology, 115*(5), 1263–1271.

VISION

Vision is easy. All an observer has to do is open his or her eyes and look around. Doing this reveals visual qualities in the environment such as shapes, light, colors, objects at specific locations, spatial layout, movement, and surface textures. This gamut of experience is created by two small detectors (eyes) connected to a powerful computer (the brain), which combine to transform the light reflected from objects first into images on the retina, then into activity in the brain, and finally into the perceptions we experience.

The transformation of environmental stimuli into an image on the retina creates a situation that could be interpreted to indicate that the mechanism for vision is simple. According to this interpretation, if stimuli are represented as a picture on the retina, the information necessary for perceiving these stimuli simply need to be extracted from this picture.

This illusion of simplicity misled early computer scientists to propose, in the 1950s and 1960s, that it would take only 10 or 15 years to solve the "vision problem"—to create a device that would be able to identify objects and navigate through the environment. These proposals turned out to be overly optimistic because although progress has been made toward designing visual pattern recognition devices that can identify faces and robotic vehicles that can drive on city streets, these devices have only a fraction of the capabilities of human vision. For example, current computer–vision programs can distinguish among a hundred or more visual categories, but only after some training, and if the objects are present in isolation. Performance becomes worse when the objects appear in real-world scenes with other objects. In contrast, humans can rapidly recognize approximately 30,000 categories of objects, and performance remains high even when the objects are presented in complex scenes.

One problem facing computer–vision devices is that the stimulus on the retina is ambiguous because any image on the retina can be created by an infinite number of objects. For example, the circular shape created by a clock viewed from across a room could also be created by a quarter viewed at arms' length or an elliptical shape tilted at an angle.

How does the visual system create effortless perceptions even in the face of this ambiguity? Researchers have approached this question phenomenologically, by having observers describe what they see; psychophysically, by measuring quantitative relationships between stimuli and perception; and physiologically, by investigating biological mechanisms responsible for perception.

The dozens of entries related to vision in this book (For example, see entries on attention; color; eye movements; impossible figures; light; motion; object perception; perceptual development; pictorial perception and art; spatial layout; vision, etc.) reflect the broad scope of research in visual perception. The purpose of this entry is to identify and briefly introduce some of the approaches and research results that are described in more detail in these other entries.

Physiologically Based Research

Since the 1950s, when recording signals from single neurons began in earnest, research on the physiological mechanisms of visual perception has grown steadily; today it is one of the dominant approaches to the study of visual perception. Recent advances that have fueled this increase in physiological research are the development of brain scanning technology, which has made it possible to measure brain activity in humans, and the increased sophistication of neuropsychological research, which studies the behavior of patients with brain damage.

The Retina

The light reflected from objects is focused by the eye's optical system to form a sharp image on the retina, the light-sensitive network of neurons that lines the back of the eye. Receptors in the retina transform the light energy from this image into electrical signals, which are then transmitted to other neurons in the retina and then leave the eye in fibers of the optic nerve. The properties of the light-sensitive pigments contained in the receptors and the wiring through which electrical signals travel determine visual functions such as spectral sensitivity (the eye's sensitivity to specific wavelengths in the visible spectrum), adaptation (adjustments of sensitivity to changes in illumination), visual acuity (detail vision), and color vision.

Destinations Past the Retina

Most of the signals leaving the eye in the optic nerve travel to the lateral geniculate nucleus in the thalamus, and about 10% travel to the superior colliculus, a structure in the midbrain that is involved in directing eye movements. From the lateral geniculate nucleus, signals are transmitted to the primary receiving area of the cortex in the occipital lobe.

Early research in vision assumed that most of the processing of visual information occurred in the occipital lobe. However, research over the last 40 years has demonstrated that visual stimuli can potentially activate over half of the cortex. The idea that visual processing occurs outside the occipital cortex was fueled by the discovery, in the 1980s, of two visual pathways, the ventral pathway, extending from the occipital lobe to the temporal lobe, and the dorsal pathway, extending from the occipital lobe to the parietal lobe. What is noteworthy about these pathways is that they are involved in different aspects of visual functioning. The ventral pathway is important for identifying and recognizing objects, and the dorsal pathway for locating objects and taking action toward them. Signals related to vision also travel to many other areas of the cortex, including the frontal lobe (important for higher-level visual functioning, such as decision processes that involve holding objects in memory during a delay, and for coordinating visual signals with inputs from other senses), and subcortical areas such as the amygdala (important for emotional behavior, it plays a role in determining the emotions expressed by facial expressions), and the hippocampus (important for memory, it contains neurons that fire preferentially to stimuli and classes of stimuli that have been encountered in the past).

Modularity of Function

There is abundant evidence that specific areas of the cortex are specialized to process information about specific types of visual stimuli. Prominent examples are the infrotemporal cortex in the ventral stream (perceiving objects), the mediotemporal cortex in the dorsal stream (perceiving motion), the fusiform face area in the fusiform gyrus on the underside of the brain (perceiving faces), and the parahippocampal place area (perceiving buildings and locations). Although these areas have been shown to be specialized for perceiving information associated with specific object properties, it is also clear that specific stimuli cause widespread activation of the cortex. For example, after a person has identified a face, he or she may have an emotional reaction to the face based on facial expressions or past experiences with the face's owner. He may also notice whether a person is looking straight at him or off to the side, and may be thinking about how attractive (or unattractive) the face is. Each of these reactions to faces has been linked to activity in different areas of the brain.

Experience and Neural Function

The question of how brain areas become specialized to process information about specific types of stimuli has been answered in two ways. One idea is that selectivity is a result of evolution and is therefore prewired into the brain. Another idea is that specialization develops due to interactions with the environment. A central mechanism responsible for this effect of experience is *experience-dependent plasticity*—the modification of a neuron's response properties that occurs due to experience, with these modifications usually causing the neuron to respond best to frequently encountered stimuli. According to this idea, neurons could become shaped to respond selectively to faces, because of continued exposure to faces, or to respond to lines with a particular orientation by exposure to lines with that orientation. The actual cause of neural specialization is likely to involve both mechanisms, although the relative importance of evolution and experience-based plasticity and the nature of their interaction are topics of debate in the research literature.

Beyond Basic Visual Qualities

One of the major trends in research on visual perception has been consideration of functions beyond basic visual qualities such as form, depth, size, color and movement. One important topic of research is the interaction of vision and action. This research has its roots in the discovery of the dorsal stream, which processes information related to locating objects and taking action relative to them, and in the suggestion that a basic goal of visual processing is not to create a conscious perception or "picture" of the environment, but to help animals control navigation, catch prey, avoid obstacles, and detect predators—all crucial functions for survival. A large amount of physiological research influenced by this line of thinking has demonstrated close links between perceiving and taking action.

A topic of research related to action is the study of how observers carry out specific actions, such as grasping an object, and one of the most intriguing areas of research involves neurons called mirror neurons. Mirror neurons, initially discovered in monkeys, fire when a monkey observes an action being carried out by the experimenter (for example, picking up a peanut), and also fire when the monkey carries out the same action. There is evidence that similar neurons exist in humans, and there is a lively debate among researchers as to what the function of mirror neurons might be, some suggesting that mirror neurons are important for imitative behaviors, determining other people's intentions, and guiding social functioning, and others taking a more cautious "wait and see" attitude regarding the functional significance of these neurons.

Behaviorally Based Research and Theories

The behavioral approach, which involves phenomenological and psychophysical investigations, was dominant in the early history of visual perception research, starting in the late 19th century. It is still an essential approach because of the importance of defining and measuring the capacities of perceptual systems, and determining the nature of the stimuli that control perception. The behavioral approach has also played an important role in motivating prominent theoretical approaches to perception.

Operating Characteristics of the Visual System

An important aspect of research in every sensory system is determining the basic operating characteristics of the system. For vision, this has involved determining functions, such as: (a) thresholds—the absolute sensitivity to light and the ability to distinguish between two intensities; (b) light and dark adaptation—the time course of the sensitivity changes that occur in response to changes in levels of illumination; (c) color vision—the perception of the colors of single wavelengths, broadband stimuli, and mixtures of wavelengths; (d) motion perception—detecting the speed and the direction of motion; (e) depth perception—judging both absolute and relative distances; (f) layout perception—perceiving the spatial layouts of environmental scenes; and (g) attention—directing attention to different places in visual displays and scenes.

This behavioral research is important for two reasons. First, it is important to define characteristics of the system and to define the perceptual phenomena that need to be explained. For example, any theory of color vision needs to explain why it is possible to perceptually match the color of any wavelength in the spectrum by mixing together three other wavelengths. The second reason behavioral research is important is that behaviorally determined system-operating characteristics can often be used to infer underlying physiological mechanisms. For example, the idea that color vision is based on three wavelength-sensitive mechanisms was proposed in the 19th century, based on color-matching data. It wasn't until almost 100 years later that physiologically based research determined the identity of the three light-sensitive visual pigments that underlie these three mechanisms.

The Gestalt Approach to Perception

The Gestalt approach, founded in the 1920s by Max Wertheimer, proposed principles to explain perceptual organization (how individual elements are grouped into larger units) and perceptual segregation (how objects are differentiated from one another, as when the buildings in the skyline are seen as separate overlapping buildings). These organizational principles, and the Gestalt credo, "The whole is greater than the sum of its parts," are still influential today because they emphasize both the importance of dealing with the basic ambiguity of visual stimulation and the importance of considering the effect of contextual factors on perception.

Ecological Approach to Perception

This approach, which was founded by J. J. Gibson, has as its premise that it is important to study perception under natural conditions and that research should focus on the moving observer. This emphasis on the moving observer contrasts with the mostly stationary observers who were making judgments in perception laboratories when Gibson first proposed the ecological approach. This focus on the moving observer has led to the identification of new stimuli, such as optic flow (the way stimulus elements in the environment flow past observers in response to their movement), and to the idea of "observer produced stimulation" (how stimuli for perception are produced by movement of the observer). Another of Gibson's ideas was that the functions of objects, which he called "affordances," (for example, "it is something to sit on" for a chair) play a role in perception. These ideas were not immediately accepted when Gibson first proposed them, but many current vision researchers have embraced the ideas that movement, action, and an object's function are important aspects of the study of perception.

Construction/Inference/Environmental Regularities

The fact that there is not adequate information in the retinal image to precisely define what is "out there" in the environment has led to the idea that the observer plays an active role in constructing perceptions from the stimulation available. This construction has been described as (a) an inferential process in which the observer infers, often unconsciously, the stimulus that would most likely have caused the retinal image (Hermann von Helmholtz's theory of unconscious inference); (b) a process that is similar to thinking or problem solving (Irwin Rock, *The Logic of Perception*); and (c) the outcome of a statistical process called Bayesian inference, in which observers take into account probabilities of occurrence of stimuli in the environment. This third approach includes the idea that observers have internalized statistical regularities in the environment—features or elements that are likely to occur in visual scenes. Examples of statistical regularities are that a blue region at the top of an outdoor scene is likely to be sky or that city scenes tend to be made up of many vertical facades.

Perceiving Complex Stimuli

Early research in vision used simple stimuli such as circular fields of light. Although the Gestalt psychologists introduced more complex stimuli, such as patterns and figure-ground displays ("figures" situated on "backgrounds"), the use of simple stimuli dominated early research in visual perception. However, current research in visual perception is concerned with how observers perceive more complex and often ecologically relevant stimuli, such as objects and environmental scenes. These stimuli have usually been presented as pictures on a computer screen, although virtual reality displays and stimuli in the actual environment have also been used. This research is related to the idea of regularities in the environment and is also associated with physiological research on higher levels of the visual system, where neurons have been found that respond to complex stimuli and that are influenced by the context of a scene.

Intersections With Cognition

The idea described earlier, that perception may involve construction, inference, and consideration of regularities in the environment, reflects the operation of cognitive processes in perception. The involvement of cognitive processes is often conceptualized in terms of top-down processing, where top-down processing refers to processing that takes as its starting point knowledge that the observer brings to a situation. An example of this knowledge is what observers have learned about the types of objects that are likely to appear in different contexts. For example, under degraded viewing conditions, as when images of objects are blurred, identification is aided by the presence of context. Thus, a blurry image depicted on a tool bench might be perceived as an electric drill, but the same image seen in the context of a bathroom counter would be perceived as a hair dryer.

Visual perception is also closely linked to cognitive processes, such as attention (the way observers scan a scene can be influenced by the observer's past experiences in observing environmental scenes), memory (memory can be enhanced for experiences rich in visual detail, and perception can be influenced by past experiences); thinking (perception can both aid thinking and, as articulated by the constructivist and other approaches, involves processes similar to those involved in thinking), and emotion (visual stimuli can create both positive and negative emotions, and emotional state can influence perceptual behavior).

From this brief survey of vision research it is clear that research in vision occurs on a number of levels. The rich tradition of psychophysical and physiological research on visual processes, such as light and color perception and the perception of depth, movement, and object properties, continues to add to knowledge about basic visual mechanisms. In addition, vision research has also become concerned with studying the perception of complex, environmentally relevant stimuli; with how observers interact with these stimuli; and with physiological processes that extend far beyond the primary visual receiving area of the cortex.

E. Bruce Goldstein

See also Action and Vision; Bayesian Approach; Computer Vision; Contrast Perception; Direct Perception; Ecological Approach; Face Perception; Gestalt Approach; Motion Perception; Navigation Through Spatial Layout; Object Perception; Perceptual Development: Visual Acuity; Perceptual Organization: Vision; Physiological Approach; Psychophysical Approach; Reaching and Grasping; Receptive Fields; Retinal Anatomy; Spatial Layout Perception, Neural; Spatial Layout Perception, Psychophysical; Statistical Learning; Top-Down and Bottom-Up Processing; Theoretical Approaches; Vision: Cognitive Influences; Visual Acuity; Visual Categorization: Physiological Mechanisms; Visual Processing: Extrastriate Cortex; Visual Processing: Primary Visual Cortex; Visual Processing: Retinal; Visual Processing: Subcortical Mechanisms for Gaze Control; Visual Receptors and Transduction; Visual Scene Perception; Visual Scene Statistics; Visual System Structure

Further Readings

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*(8), 617–629.

Chalupa, L. M., & Werner, J. S. (Eds.). (2003). *The visual neurosciences.* Cambridge: MIT Press.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191–197.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Goldstein, E. B. (1998). When does visual processing become cognitive? *Contemporary Psychology, 43,* 127–129. (Review of S. M. Kosslyn & D. N. Osherson, *An invitation to cognitive science, Vol. 2, Visual cognition*)

Goldstein, E. B. (2001). Cross-talk between psychophysics and physiology in the study of perception. In E. B. Goldstein (Ed.), *Blackwell handbook of perception* (pp. 1–23). Oxford, UK: Blackwell.

Goldstein, E. B. (2010). *Sensation and perception* (8th ed.). Belmont, CA: Wadsworth Cengage Learning.

Goldstein, E. B. (in press). *Cognitive psychology* (3rd ed.). Belmont, CA: Wadsworth Cengage Learning.

Helson, H. (1933). The fundamental propositions of Gestalt psychology. *Psychological Review, 40*(1), 13–32.

Hubel, D. H. (1982). Exploration of the primary visual cortex, 1955–1978. *Nature, 299*(5883), 515–524.

Ishai, A. (2008). Let's face it: It's a cortical network. *Neuroimage, 40*(2), 415–419.

Metzger, W. (2006). *Laws of seeing* (L. Spillman, Trans.). Cambridge: MIT Press. (Original work, *Gesetze des Schens,* published 1936)

Milner, A. D., & Goodale, M. A. (2006). *The visual brain in action* (2nd ed.). New York: Oxford University Press.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences, 11*(12), 520–527.

Rock, I. (1983). *The logic of perception.* Cambridge: MIT Press.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–580). Cambridge: MIT Press.

Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences, 10*(7), 301–308.

VISION: COGNITIVE INFLUENCES

To be useful in guiding behavior, perceptual systems must derive representations of the world that allow for fast, accurate responses to the world. Perceptual systems draw on various processes to optimize behavior. For example, early feature detection processes in vision can detect co-occurrences between features, and detecting these co-occurrences permits rapid grouping of features into longer contours. But, cognitive influences also exist in vision, as when our expectancies or prior experience influence vision. These cognitive influences also can optimize behavior by altering visual representations.

There are many empirical demonstrations of cognitive influences in vision, and several examples are discussed in this entry. However, more important than demonstrations of cognitive influences are how these cognitive influences operate and affect vision. Does cognition influence vision only *after* visual perceptual processes have completed, as a way to "clean up" perception? Or, does cognition influence these perceptual processes *as they operate,* altering the operation and outcome of perception directly?

The foregoing questions point out that a still-standing issue in visual perception is the manner in which cognition influences perception. There are two dominant theoretical approaches to understanding the role of cognition in vision. One view, the bottom-up account, assumes that vision proceeds in a bottom-up or feedforward manner in which higher-order percepts are created from combinations of lower-level features. Strong versions of this bottom-up account propose that early visual processes are unaffected by—that is, encapsulated from—cognition and other later-level processes, as Zenon Pylyshyn, among others, has argued. In short, observers' expectations should not affect what they perceive. Under such a view, cognitive influences in vision are postperceptual and occur after visual processes have completed.

In contrast, the other view, the interactive account, proposes that vision operates in an interactive manner, in which both bottom-up information from the external world and top-down information (e.g., an observer's goals, expectancies, or prior experience) combine to determine visual perception. According to this interactive account, top-down information can be used to guide perception in the face of incomplete bottom-up information, and this top-down influence operates on perceptual processes themselves. In this way, vision can be viewed as performing what Hermann von Helmholtz termed "unconscious inferences" or unconscious problem solving.

The literature is full of arguments for both bottom-up and interactive accounts of various visual processes. The persuasiveness of these arguments rests on a number of methodological issues that arise when distinguishing bottom-up and interactive accounts. In short, it can be difficult to understand the mechanism that produces a cognitive influence in vision with psychophysical data because many experiments can be interpreted in terms of both bottom-up and interactive accounts. Given this, an instructive approach is to focus on how these cognitive influences shape vision. To understand how cognition affects vision—whether in a bottom-up or an interactive manner—requires an understanding of the methods used to investigate cognitive influences and how the various mechanisms can be disentangled. Many of the relevant methodological issues appear in early studies of cognitive influences in vision, including some of the earliest studies on figure-ground perception.

A Historical Example

Figure-ground perception involves the visual system assigning some regions as foreground figures and others as backgrounds. The earliest work on figure-ground perception arose from the Gestalt psychologist Edgar Rubin, who outlined a number of bottom-up cues that could be used to solve figure-ground problems. For example, Rubin noted that small, surrounded regions are more likely to be perceived as figures than as grounds. The same holds for convex regions and for symmetric regions. These Gestalt "laws" of figure-ground perception thus represent bottom-up cues that the visual system can use in determining figure-ground relations.

But, Rubin also reported an effect of prior experience in figure-ground perception in which previous exposure to a figure-ground stimulus altered figure-ground perception when the display was viewed again at a later time. This finding suggested a cognitive influence in figure-ground perception, namely, that previous experience influences which area is perceived as figure and which as ground. Rubin trained observers to have specific figure-ground interpretations for different displays. For example, in the figure-ground display in Figure 1, an observer might be trained to perceive the small white region as figure. Rubin reported that training observers to see a region as figure increased the tendency for that region to be perceived as figure at a later time. Observers tended to interpret displays in the manner they had been interpreted previously, during training.

Figure 1 Figure-Ground Display

On the surface, Rubin's findings appear to support a role for prior experience and familiarity on visual perception that could operate in a top-down manner. The trained figure becomes a familiar shape that guides later figure-ground perception. Unfortunately, Rubin's methods do not permit an unequivocal demonstration of a top-down effect of familiarity. For example, a response bias could explain the results, in which observers choose as the figure the more familiar of two regions in a figure-ground display; figure-ground processes might not be affected by familiarity under such an account. Rubin asked observers to provide explicit reports of figure-ground perception (i.e., which region is figure?), and such explicit reports could result from visual processes different from those processes that determine figure-ground perception. An explicit figural report could be determined from figure-ground processes themselves, but it could also be determined from later recognition processes, in which a familiar shape is more likely to be chosen as figure over a less familiar region.

A bottom-up account of the prior experience effect is therefore quite plausible. Familiarity effects might arise because figure-ground processes extract the likely figure based on bottom-up Gestalt cues; a later process then recognizes familiar figures faster and more accurately than unfamiliar figures, allowing these familiar figures to be chosen as the likely foreground figure in the display. Thus, to argue convincingly for a cognitive, top-down effect over visual perception requires one to rule out various bottom-up alternatives.

Disentangling the Mechanisms of Cognitive Influences

As illustrated by Rubin's results, cognitive influences can be implemented with either bottom-up and top-down, or interactive, processes. One of the most salient alternative interpretations to a top-down cognitive influence in perception is that of a response bias or change in response criteria, as opposed to a change in perception itself. Consider another example from the literature: There have been numerous demonstrations showing that objects are detected more easily in semantically consistent scenes (e.g., recognizing a loaf of bread in a kitchen scene) than in semantically inconsistent scenes (recognizing a mailbox in a kitchen scene). Although these results demonstrate that cognition is influencing vision, the precise influence is unspecified. This semantic context effect might not be due to a top-down influence of scenes on object perception. Instead, under conditions of uncertainty—such as briefly exposed scenes or the necessity for speeded responses—observers can use the scene to make guesses about the objects present, such as guessing that an object in a kitchen scene was a loaf of bread and not a mailbox. Under such a guessing account, the perception of the object may not be altered by scene consistency. A guessing account supports a role for prior knowledge in scene perception, but this guessing account could operate in a bottom-up manner, after object recognition was completed.

A straightforward approach to eliminate such a guessing strategy is to have observers perform forced choice tasks in which both response alternatives were equally plausible (or equally implausible) for the context. A classic example comes from the word superiority effect, in which a letter is perceived faster and more accurately when it appeared in a word than in a nonword. In the original procedure, a word (e.g., “WORK”) or a nonword (“RWOK”) was presented briefly, and observers were asked to report which of two letters (”K” or “M”) appeared in a certain position in the word (the last position in this example). Importantly, for the word, both letters would create a meaningful word (“WORK” versus “WORM”), eliminating any advantage of guessing the letter based on the context.

When such a forced choice procedure is used to study object recognition in semantically consistent and inconsistent scenes, objects in consistent scenes are not identified faster than those in inconsistent scenes. Thus, scene consistency might not affect object recognition processes and alter object perception. Instead, scenes might establish expectations about the objects that might be present, which, in turn, establish biases in reporting that certain objects were present in a briefly presented scene. Thus, although there is a cognitive influence of scenes on object perception, this influence appears not to alter the perceptual processes themselves. Instead, the cognitive influence might reduce the uncertainty associated with identifying objects that are only partially perceived because of brief exposure durations (or, in real-world scenes, because the eyes have moved quickly to another object).

In addition to using forced-choice tasks to reduce response biases, indirect measures also can help minimize response biases by having observers perform a task that is unrelated to the cognitive processes that might affect vision. For example, in both the word superiority and scene context examples previously mentioned, observers gave explicit recognition reports (e.g., which letter was present, or which object was present?); but explicit letter (or object) recognition could be performed by recognizing an entire word or scene. Rubin's procedure to study familiarity effects also relied on an explicit report of figure-ground perception. As noted earlier, however, explicit figure-ground reports could be based on the results of a figure-ground process or a later recognition process.

The use of indirect measures is illustrated in studies of the role of object familiarity on perceptual organization and image segmentation. Is visual image segmentation influenced by object familiarity in a top-down manner? To address this question, observers performed an image segmentation task that was unrelated to the familiarity of the objects being segmented, thereby making object familiarity irrelevant to the task. Observers reported if two small Xs were on the same shape or on different shapes. The shapes themselves—whether familiar upright letters or less-familiar upside-down letters—were not required to perform the segmentation task, making it possible to rule out response biases similar to those in scene context effects discussed earlier.

However, even when response biases can be excluded, as in studies of the word superiority effect, distinguishing bottom-up and interactive accounts can be complicated by other factors. Behavioral results that suggest context influences perception can be explained with either bottom-up or interactive models, for example. The typical interactive account for the word superiority effect proposes that letters are represented at an earlier level of processing than are words; top-down feedback from the word level alters processing at the letter level, allowing letters to be recognized faster and more accurately when they appear in a word than when they appear in a nonword. But, the same behavioral result could be explained by a bottom-up model in which letter co-occurrences within words are represented relatively early and influence later word representation processes. The "__RD" in the last two positions occurs in English, but the visually similar "__PD" (as might occur in a nonword) does not occur in English. The legal co-occurrence might allow "D" to be distinguished more easily from other letters than the illegal (or infrequent) co-occurrence. In short, familiarity—the information that allows words to be distinguished from nonwords—could be represented relatively early in the system and could produce behavioral effects in a bottom-up manner, possibly through faster bottom-up processing for legal or common co-occurrences than for illegal or uncommon co-occurrences. Such early knowledge would indicate that prior experience need not be an entirely cognitive influence. Instead, learning perceptual regularities—such as the regularities of English orthography—might impact perceptual systems as well as cognitive systems.

Finally, in addition to faster bottom-up processing accounts, some approaches to studying cognitive influences on perception propose that targets and context can be combined independently, without interactive processing. Gregg Oden and Dominic Massaro's fuzzy logical model of perception performs such a combination and explains the word superiority effect without appealing to top-down influences on letter perception. The fuzzy logical model is an independence model in which different sources of information are used to determine letter identity and word identity. The information used to determine word identity does not influence information used for letter identification. The most probable letter and word identities are computed from the perceptual information, and these probable identities then are combined to determine the most probable interpretation of the entire stimulus (letter and word).

This combination occurs at a later integration stage, not at a perceptual evaluation stage; the perception of the letter information itself is not altered by the surrounding word context. Studies that orthogonally manipulate letter shape information and word context show that these factors combine additively, consistent with the independent integration in the fuzzy logical model.

A Recent Example: Cognitive Influences on Visual Completion

To highlight how the foregoing methodological issues can be used to support an interactive cognitive influence in visual perception, this entry concludes with an example from visual completion. In complex visual scenes, objects are often partially occluded by other, closer objects, as when a car is parked behind a lamppost and the middle section of the car is visually obstructed by the post. The visual system completes the occluded regions of objects, so that an occluded object is perceived as continuing behind the occluder. The car behind the post is perceived as just that—a car—and is not perceived as two separate and accidentally aligned ends.

This visual completion process is thought to occur relatively early in vision (i.e., is preattentive), and most accounts of visual completion argue that this is a bottom-up visual process. These accounts argue that completion is performed on the basis of image information, such as the alignment of an object's edges on either side of the occluder. Cognitive factors, such as object familiarity or expectancy, do not directly affect completion under these bottom-up accounts. Indeed, Gaetano Kanizsa showed that expectations do not affect visual completion. As shown in Figure 2, a row of unoccluded objects (Figure 2a) cannot override the tendency to complete occluded objects (Figure 2b). Bottom-up completion occurs even when the occluded object—an abnormally long horse—violates cognitive expectations and experience.

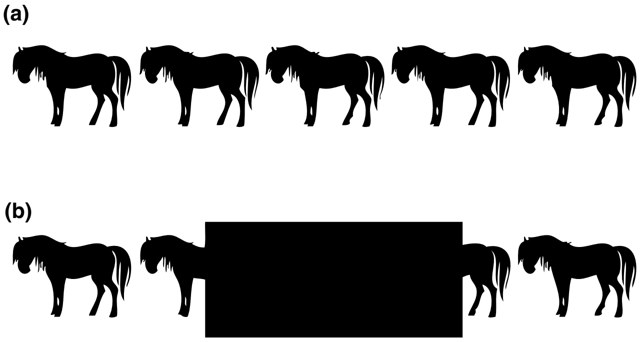


Figure 2 Visual Completion

Note: (a) A row of unoccluded objects cannot override the tendency to complete (b) Occluded objects.

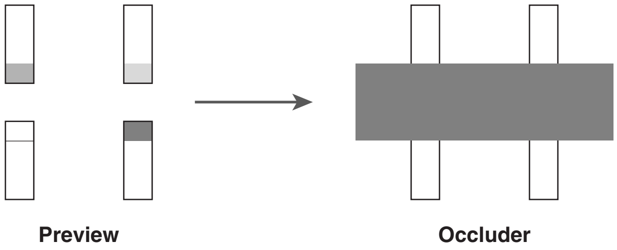


Figure 3 Influence of Short-Term Memory on Visual Completion

Notes: Observers viewed four rectangles, each having a small color patch at one end. An occluder then appeared, allowing the aligned visible ends to complete into two long rectangles.

In response to bottom-up accounts of visual completion, Hyunkyu Lee and Shaun Vecera noted that observers in completion tasks typically never need to actively remember information behind the occluder. Because the occluded visual information does not need to be remembered, bottom-up cues may tend to dominate completion. To ask if visual short-term memory could influence visual completion, observers remembered visual information that was later occluded. Observers viewed four rectangles, each having a small color patch at one end (Figure 3). An occluder then appeared, allowing the aligned visible ends to complete into two long rectangles, similar to Kanizsa's demonstration (Figure 2).

When observers were not remembering the small color patches, the visible rectangle ends were perceived as two longer rectangles, despite observers' knowledge that four rectangles had been present in the display. However, when observers were instructed to remember the color patches, the visible rectangle ends were no longer completed. Observers did not perceive the aligned visible ends of the rectangles as being completed behind the occluder; instead, observers interpreted the display as containing four separate objects. These results suggest that visual short-term memory can influence—and, more important, override—visual completion processes in a top-down manner. The use of an implicit task that was unrelated to completion eliminated various response bias interpretations. And, demonstrating that completion was actually halted (as opposed to merely slowed) is problematic for the various bottom-up accounts discussed earlier.

Final Thoughts

There are numerous demonstrations of cognitive influences on visual perception, ranging from phenomena such as the word superiority effect to scene context effects on object recognition. The mechanisms that produce these cognitive influences—feedforward (bottom-up) and interactive (top-down) accounts—are less understood. Further, most demonstrations of cognitive influences can be explained with either account. Understanding the various implementations of these accounts and disentangling them is more than an academic exercise in experimental design; instead, using correct methods is critical for moving beyond a mere demonstration of a cognitive influence to an understanding of how such an influence affects vision.

Shaun P. Vecera and Hyunkyu Lee

See also Amodal Perception; Modularity; Perceptual Organization: Vision; Top-Down and Bottom-Up Processing; Visual Filling In and Completion; Visual Scene Perception

Further Readings

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 213–253). Hillsdale, NJ: Erlbaum.

Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General, 127*(8), 398–415.

Lee, H., & Vecera, S. P. (2005). Visual cognition influences early vision: The role of visual short-term memory in amodal completion. *Psychological Science, 16*(10), 763–768.

Oden, G. C., & Massaro, D. W. (1978). Integration of featural information in speech perception. *Psychological Review, 85*(3), 172–191.

Pylyshyn, Z. W. (1999). Is vision continuous with cognition? The case of impenetrability of visual perception. *Behavioral and Brain Sciences, 22*(3), 341–423.

VISION: DEVELOPMENTAL DISORDERS

The development of human visual behavior involves many interleaved processes. These processes have been called *sensation, perception, cognition, attention,* and *action,* depending on the interests of the investigator. However, it is often impossible to draw boundaries between them, and their development is highly interlinked. Many researchers take the standpoint that any visual act involves all these processes, in cascading sequences and re-entrant loops of neural activity through multiple brain areas and networks. Neither in normal development nor in disorders of development can they be usefully considered in isolation. In specific disorders, sometimes it may be possible to localize the original source of the pathology, but even then the developmental consequences of this pathology may spread through many levels of processing. It is also important to recognize that many pathologies, both genetic syndromes and brain damage, have far-reaching effects and may impair other aspects of function (social, linguistic, etc.) in addition to their perceptual/cognitive effects.

In some developmental disorders, there is a clear genetic origin (e.g., a defective protein critical for photoreceptor function, or a known gene or genes expressed in the brain). In others there is some kind of early accident or injury that has developmental consequences. This entry discusses examples of both genetic and environmental effects operating at levels from the eye to all levels of the cortex. However, in many disorders, a genetic or environmental cause may be hypothesized but cannot be pinpointed.

This entry concentrates on disorders of visual processing, and their links with cognitive, attention, and motor function. The other senses, particularly hearing, also show developmental pathologies analogous those of vision. The developmental consequences of hearing disorders are above all in the realm of language development, which is beyond the present scope.

Outline of Normal Development

Developmental disorders must be seen in the context of the process of normal perceptual development. This has several key aspects:

- Visual acuity, studied with behavioral and electrophysiological methods, shows a development from less than 1/10th of adult values at birth, with rapid development over the first few months, but with adult values not attained until 4 to 6 years of age. This is limited in part by the immaturity of photoreceptors, but development of neural systems in the eye and brain is undoubtedly necessary.

- The visual abilities of the newborn are primarily mediated by subcortical systems. The characteristic selectivity of the primary visual cortex emerges over the first 3 to 6 months of life, first for contour orientation and later for direction of motion. These two forms of selectivity can be regarded as the basis of subsequent processing in the ventral cortical stream (occipital to temporal, for identification) and the dorsal stream (occipital to parietal, for spatial actions), respectively.

- Binocularity, the combination of information from the two eyes on the same cortical neurons, also develops rapidly from about 3 months, rather later than other functions of the primary visual cortex.

- Facelike stimuli are a target of infants' fixation from birth, but the development of specialized cortical systems for face processing is a long process and is not mature before middle childhood.

- Global processing in the extrastriate visual cortex is revealed by the integration of information in patterns of motion (for dorsal-stream areas) and static form (for ventral-stream areas). This continues on from initial visual cortical development. Global motion sensitivity appears in infants more rapidly than for global form; however, in later development through childhood, the dorsal stream function of global motion takes longer to reach adult levels than ventral form processing, and as discussed later it is more vulnerable.

- Apart from eye movements, visually controlled action systems (reaching, grasping, locomotion) develop from about 5 months onward. Progressive integration of visual action planning and the control of visual attention required for this also develop through infancy. These are also predominantly dorsal functions, and again are particularly vulnerable. Attention consists of multiple developmentally dissociated systems. Earlier maturity in selective attention (before 6–7 years), for example, contrasts with the continued development of sustained attention into adolescence and rapid development of executive attention between 7 and 11 years.

Refraction and Focusing

The first stage in vision is the optics of eye bringing images to a sharp focus on the retina. This depends on the size and shape of the eyeball. In young infants, the eye is on average hyperopic or longsighted—but the action of accommodation (focusing) adjusts the lens to overcome this, so that in the first few months, the infant's eyes are usually focused on near objects, although they will adjust in the appropriate direction for an object that moves toward or away from the child.

The refraction of the eyes results from a developmental process: It is known from animal models that a blurred image, resulting from refractive errors present early in life, serves as a stimulus to modulate the growth of the eye so that the blur is reduced, a process known as *emmetropization* (normalization). It is not known why this process sometimes fails to correct the initial hyperopia. By age 9 months, about 5% of infants show significant refractive errors where the eye is poorly focused—in Caucasian populations mostly excessive hyperopia, often associated with marked astigmatism (i.e., lines at different angles/slant/orientation cannot be sharply focused equally at the same time). As well as reducing the quality of spatial information available to the infant, these conditions have longer term consequences. They are associated with amblyopia and strabismus (discussed later), developments that may be avoidable if the refractive errors are corrected with appropriate spectacles prescribed and worn from infancy (and that do not appear to block the natural process of emmetropization). These infant refractive errors are also associated with a broader range of subtle delays in development of visual attention and in perceptual, cognitive, visuomotor, and spatial abilities, first identifiable in the second year of life and persisting into school entry, perhaps associated with development of frontoparietal systems for spatial cognition and attention. However, it is not known how far these problems share a common neurodevelopmental origin with hyperopia or how far they result from refractive degradation of the visual input.

Myopic (shortsighted) refractive errors are rare in infancy, but increasingly common from middle childhood onward. The incidence varies considerably between different ethnic groups, with higher incidence of myopia in infants of Asian ethnicity. Again, it is not known why the natural corrective process fails, but it is clear that there are developmental effects, in that the persistent accommodation associated with visual near work has sometimes been found to exacerbate the development of myopia.

Retinal Disorders

The retina, which transforms light into neural signals, is a significant location for developmental pathology:

- In a number of genetic disorders, the photoreceptors degenerate progressively from birth.
- Retinopathy of prematurity is a form of retinal damage caused by excessive oxygen that may be delivered in the intensive care of very premature infants. This is now well recognized where neonatal care is most advanced, but remains a problem in some countries and in cases where the survival of extremely premature infants has to be balanced against the risk of ocular effects.
- Atrophy of the optic nerve (the fibers of the retinal ganglion cells) is a condition that may be associated with other developmental defects.

These are all significant causes of visual impairment and blindness of great concern in pediatric ophthalmology. However, this entry concentrates on conditions associated more with central processes of perception. In fact, as other causes of child blindness become better managed, problems due to central cerebral disorders increasingly dominate the epidemiology of childhood visual disability in developed countries.

Binocularity, Amblyopia, and Strabismus

One of the most common areas of visual developmental disorder is in the development of binocularity. The cortical connections that enable combination of information from the two eyes first become functional around three to four months of age. However, effective binocular vision depends on a coordination of this sensorineural process with precise motor alignment of the eyes. If the eyes are misaligned, signals from corresponding points of left- and right-eye images do not come together in the cortex, the correlated signals needed to maintain connections from the two eyes to the same cortical cell are absent, and these binocular connections break down. Conversely, fine alignment of the eyes depends on the ability to detect matches and mismatches of the two eyes' images and adjust the convergence of the eyes accordingly. It follows that there is a feedback loop maintaining binocular vision that is easily broken, either by a defect on the sensory side (as in albino individuals where some optic nerve fibers are misrouted and do not reach the correct side of the brain) or on the motor side (for example, if one or more eye muscles are weak or paralyzed). This breakdown appears as a misalignment of the eyes known as *strabismus* (often called *squint* in the United Kingdom), in which the failure of binocular cortical mechanisms leads to *stereoblindness*—the absence of stereoscopic depth perception. The convergence of the eyes is closely linked with the accommodation of the lens, which may explain why childhood hyperopia, in which abnormal degrees of accommodation are required, is often associated with the development of strabismus.

Loss of stereopsis impairs some visuomotor skills (try playing table tennis with one eye closed!) and is taken as a disqualification for some occupations (e.g., military pilot). However, the diversity of cues for depth perception makes it possible to compensate for this loss in many situations; for example there are a number of examples of successful sportsmen in high-speed ball games without binocular vision. The social difficulty of maintaining eye contact with uncorrected strabismus probably has greater impact on everyday life.

The readiness with which the developmental feedback loop between eye movements and cortical connections can be broken may explain why binocularity is vulnerable in infancy and strabismus is frequent in all kinds of neurodevelopmental disorders (Down syndrome, prematurity, perinatal brain injury, etc.).

Animal models and clinical experience indicate that the establishment of binocularity is associated with a critical period. If the coordinated signals are not provided to the cortex before about three months of age in kittens, and probably about two years in humans, binocular connections in the brain cannot be re-established, even when the eyes are subsequently aligned. This limited period of plasticity in brain connections is also seen in the conditions known as *amblyopia.*

Amblyopia is defined as a reduction of visual acuity, usually in one eye, that has no detected organic cause in the eye and cannot be improved with spectacle correction. It is believed to be a developmental disorder of neural connectivity in the visual cortex when one eye's image is degraded in some way. During a critical period when synaptic mechanisms are plastic, there is a competition between the two eyes for cortical connections, which are diminished for the eye that is providing degraded information. In *deprivation amblyopia,* the most straightforward case, one eye has pattern vision reduced or abolished (e.g., by a dense cataract). Poor acuity remains even after the obstruction of vision is removed. *Anisometropic amblyopia* occurs when one eye has a greater refractive error (usually hyperopia) leading to greater image blur than the other eye. This acts as a partial form of deprivation amblyopia, where the relative deprivation is for fine detail rather than all pattern vision. *Strabismic amblyopia* is reduced acuity in the eye that is misaligned, and so not used for fixation, in strabismus. All these conditions become less and less easy to correct for longer periods of deprivation, and so it is important for any clinical intervention (e.g., to correct anisometropic focus) to be as early as possible. However, the critical period for acuity development is not necessarily the same as that for binocularity. It is also important to understand that the critical period for disruption of a visual function is not necessarily the same as the period during which that function is established in normal development.

Although amblyopia is commonly assessed in terms of visual acuity, the actual visual consequences of visual deprivation are considerably more complex. Amblyopes lose not just the ability to detect fine detail but also some of its spatial organization, so that they complain that patterns may appear "scrambled" and suffer from "crowding" effects when trying to read fine print. It appears that cortical input from the deprived eye is not just reduced but may also be disorganized. Studies of children who were surgically treated for congenital cataract suggest that there are important aspects of plasticity in central visual processing that are not captured by measures of binocularity and acuity. Children with only a few months of visual deprivation before a cataract was removed show persisting impairment in recognizing faces and in tasks requiring global form and motion processing. These effects are much greater than would be expected from any persisting acuity loss. They occur even though at the age vision was restored, these aspects of high-level visual processing were quite immature, and have been referred to as sleeper effects for this reason. These findings indicate that early visual experience is required to set up the infrastructure for later development involving cortical processing in both the dorsal and ventral streams.

Cerebral Visual Impairment

The period around birth is one of particular hazard for brain development: Difficulties in the birth process can lead to damaging episodes of ischemia (reduced blood supply) and hypoxia (oxygen deprivation). Very premature birth is often associated with damage to developing white matter (nerve fiber tracts), which in turn leads to impaired cortical development. Early infections (e.g., viral meningitis) present another hazard. *Cerebral visual impairment (CVI)* refers to a deficit of visual behavior as a result of such brain damage, typically identified in infancy by poor fixation and following, and the failure to develop visually guided reaching. It is unlikely to be an isolated impairment; the underlying neurological damage can commonly lead to cerebral palsy, developmental delay, and/or other sensory impairments. It is important to discover the child's visual capabilities, but these should be considered as part of an overall pattern of capability and disability for the individual child.

The term *delayed visual maturation* can be a source of confusion. It may be applied when children present with visual inattention in the first months of life, which may show some long-term improvement, but generally indicates an enduring deficit of CVI. However, a few children appear visually inattentive in the first months, without known neurological damage, and then recover normal visual behavior by six months or soon after. This pattern suggests a delay in the onset of development of cortical visual mechanisms, but the reasons for such a delay are not understood.

Cerebral visual impairment should not be considered an all-or-none phenomenon. The brain basis of visual processing is complex, and perinatal brain injury can lead to a range of deficits from profound loss to more subtle impairments of specific function. Follow-ups of children born very prematurely or who suffered early insults to the brain have shown a range of deficits even in those who would not be considered clinically visually impaired. These usually appear most marked in tasks involving dorsal stream function—for example, global motion perception, biological motion detection, visuo-motor skills; they also show up in tests of visual selective and sustained attention and executive control.

As in adults, differences in the location of brain damage may affect the outcome. For example, perinatal right hemisphere damage has been found to affect the perception of global organization, while left hemisphere damage has more effect on analysis of local features. However, these deficits are less severe and less distinct than the effects of corresponding lesions in adults, reflecting the plasticity whereby one hemisphere may be able to take over function from another during development.

Developmental Syndromes

Several developmental syndromes, of known or hypothesized genetic origin, have widespread behavioral and cognitive effects but include characteristic disorders of perception and related functions. One of the most striking is Williams syndrome (WS), a rare disorder resulting from the deletion of 20 to 30 genes on one copy of chromosome 7. Individuals with WS have an uneven cognitive profile; they are "hyper-social," relatively fluent in language production, generally good at face and object recognition, but severely impaired in visuo-spatial tasks such as block construction and drawing, and poor at planning and executing visuo-motor tasks (e.g., using tools and implements, stair descent, walking over uneven surfaces). This pattern suggests a deficit of the dorsal compared to the ventral stream, a proposal that is supported by findings of poor thresholds for global motion compared to global form patterns, difficulty in using orientation to control a visual posting action compared to perceptual matching, and poor calibration of visually guided actions. Magnetic resonance imaging (MRI) studies of WS have found structural anomalies and functional underactivation of dorsal stream (parietal lobe) structures. However, the deficits are not restricted to these levels of dorsal stream; individuals with WS have severe problems in visual-constructive tasks (block designs and drawing), and in various aspects of visuo-motor control. Their drawings are much more impaired in global configuration than in the production of local details. In common with many developmental disorders, they show problems of executive function; in WS these are disproportionately great if the task to be controlled has a spatial element, suggesting that transmission of spatial information to frontal systems within the dorsal stream is specifically disrupted. Wayfinding is also an area of difficulty; it can be argued how far this is a dorsal stream deficit and how far it implicates systems in the hippocampus, which would not usually be considered to be in the dorsal stream. Neuroimaging studies of WS have found functional and structural anomalies in dorsal stream areas in the parietal lobe (although these are by no means the only anomalies reported).

Other developmental syndromes have intriguing commonalities with, and differences from, WS. Autism and autistic spectrum disorders show heritability, which suggests a genetic origin, although the specific genes are not yet identified. Affected individuals show a complete contrast to WS in their social withdrawal and do not have a general visuo-spatial deficit, but do share the WS characteristic focus on local pattern details, to the degree that they may actually be at an advantage to typically developing children in tasks where local details have to be separated from the overall global structure, as in some embedded-figures tasks.

Fragile-X is a disorder with a clear single-gene basis, some overlap with autistic behavior patterns, and problems with spatial construction tasks somewhat analogous to those found in WS. In all three disorders, sensitivity to the global coherence of motion patterns is more affected than analogous sensitivity to global static form patterns. Thus, this sign of dorsal stream impairment is clear-cut, but it is not diagnostic of a particular disorder.

Dorsal Stream Vulnerability

The differential deficit of global motion over global form sensitivity is not limited to the genetic syndromes previously described. It is also found in hemiplegic children (who suffered perinatal brain injury), as a result of early visual deprivation by cataract, and in developmental dyslexia. This has led to the hypothesis that the dorsal stream shows greater vulnerability than the ventral stream in development, whether to external insults or to anomalous pathways of neural development. This idea is supported by the broad range of visuomotor and attentional deficits found in mild cerebral impairment. Some researchers propose that the differential effects originate in the magnocellular cells of the pathway to the visual cortex, relative to parvocellular cells. However, the underlying reason for dorsal stream vulnerability is not yet known.

Disorders of Attention

Many levels of visual processing control, or are modulated by, visual attention in some form, and this process appears to be highly vulnerable in cerebral visual problems. In early infancy, immature cortical function makes it difficult to shift attention from a central target if a novel stimulus appears to one side. This “sticky fixation” persists in cerebral damage, sometimes to one side if the damage is unilateral (infant hemispherectomy, in which a whole dysfunctional hemisphere is surgically removed, provides an extreme example).

Later in development, other aspects of attentional control are sensitive indicators of dysfunction. Executive control, including the ability to switch between the perceptual demands of different tasks and to inhibit prepotent responses, has been a particular focus of concern in a range of disorders. The most widespread, but not necessarily well defined, is attention deficit hyperactivity disorder (ADHD), which appears to include failures both of executive control and of sustained attention. Structural neuroimaging studies suggest that it is associated with a delay of cortical development, particularly but not only of the frontal lobes.

Many children with developmental disorders show marked attentional deficits beyond those expected from their general cognitive delays (e.g., children with WS and autism). It is possible that the parietal and frontal brain areas involved in attentional control overlap with those responsible for the dorsal stream previously discussed.

Disorders of Face Processing

In adult neurological cases, prosopagnosia is a specific disorder in which patients fail to recognize individuals’ faces without corresponding failures in other types of object recognition. It is presumed to reflect damage to specific brain areas involved in processing the facial image, such as the fusiform face area (FFA). Poor development of this area, and relatively poor ability to identify faces, has been identified as a feature of autism. However, autistic children, as part of their social disability, seem to find eye contact with other faces aversive. Therefore, it is debatable whether a deficit in the FFA is a primary feature of autism. There may be a developmental sequence in which autistics lack infants’ normal preference for looking at faces (possibly due to the way that these stimulate the amygdala), and so do not gain the experience that leads to mechanisms of face expertise becoming established in the FFA.

Some adults and children, without any autistic disorder, known brain damage, or other intellectual or visual impairment, show abnormally poor ability to identify faces. Because there is no evidence for any events that precipitate this condition, it has been labeled *developmental propsopagnosia*. The deficit may be quite striking, but its detailed pattern differs among individuals (and is different from that in children whose face recognition has been impaired by early visual deprivation).

The findings of cases of developmental prosopagnosia raise the question of whether other specific perceptual anomalies of this kind may exist in otherwise typical individuals. One example, although not usually considered a disorder, is the cross-linkage between different sensory modalities that leads some individuals to report vivid synesthesia (e.g., color sensations associated with specific words). However, neither in synesthesia nor in developmental prosopagnosia do researchers have any knowledge of the origin or the atypical course of development by which these anomalies appear.

Future Research

Overall, the occurrence of specific developmental disorders of visual processing is a promising area for understanding the relationship between perceptual abilities and the developing organization of brain systems. Advances in visual assessment offer a route for the early identification of children at risk of developmental problems, and hopefully will increasingly lead to strategies of rehabilitation for such children. However, there are many challenges; a proper understanding will only be achieved when researchers have identified not only the origins of disorders, but also the developmental sequences through which initial anomalies lead to characteristic, often widespread patterns of perceptual deficit.

Janette Atkinson and Oliver Braddick

See also Amblyopia; Cortical Reorganization Following Damage; Experience-Dependent Plasticity; Infant Perception; Low Vision; Neuropsychology of Perception; Perceptual Development: Attention; Perceptual Development: Color and Contrast; Perceptual Development: Face Perception; Perceptual Development: Hearing; Perceptual Development: Imitation; Perceptual Development: Intermodal Perception; Perceptual Development: Object Perception; Synesthesia; Visual Disorders: Blindness

Further Readings

Atkinson, J., & Braddick, O. (2003). Neurobiological models of normal and abnormal visual development. In M. De Haan & M. Johnson (Eds.), *The cognitive neuroscience of development* (pp. 43–71). New York: Psychology Press.

Atkinson, J., & Braddick, O. (2007). Visual and visuocognitive development in children born very prematurely. In C. von Hofsten & K. Rosander (Eds.), *From action to cognition: Progress in brain research* (Vol. 164, pp. 123–149). Amsterdam: Elsevier.

Atkinson, J., & Braddick, O. (2008). Vision disorders and visual impairment. In M. Haith & J. Benson (Eds.), *Encyclopedia of infant and early childhood development.* Oxford, UK: Elsevier.

Atkinson, J., Braddick, O., Anker, S., Curran, W., Andrew, R., Wattam-Bell, J., et al. (2003). Neurobiological models of visuospatial cognition in children with Williams syndrome: Measures of dorsal-stream and frontal function. *Developmental Neuropsychology, 23*(1/2), 139–172.

Atkinson, J., & Nardini, M. (2009). Visuospatial and visuomotor development. In J. Reed & J. W. Warner-Rogers (Eds.), *Child neuropsychology: Concepts, theory and practice.* Oxford, UK: Wiley-Blackwell.

Bellugi, U., Lichtenberger, L., Mills, D., Galaburda, A., Korenberg, J. R. (1999). Bridging cognition, the brain and molecular genetics: Evidence from Williams syndrome. *Trends in Neurosciences, 22*(5), 197–207.

Braddick, O., Atkinson, J., & Wattam-Bell, J. (2003). Normal and anomalous development of visual motion processing: Motion coherence and 'dorsal stream vulnerability.' *Neuropsychologia, 41*(13), 1769–1784.

Daw, N. W. (2006). *Visual development* (2nd ed). New York: Plenum Press.

Le Grand, R., Cooper, P. A., Mondloch, C. J., Lewis, T. L., Sagiv, N., de Gelder, B., et al. (2006). What aspects of face processing are impaired in developmental prosopagnosia? *Brain & Cognition, 61*(2), 139–158.

Maurer, D., Lewis, T. L., & Mondloch, C. (2005). Missing sights: Consequences for visual cognitive development. *Trends in Cognitive Science, 9*(3), 144–151.

Rubia, K. (2007). Neuro-anatomic evidence for the maturational delay hypothesis of ADHD. *Proceedings of the National Academy of Sciences USA, 104*(50), 19663–19664.

Schultz, R. T. (2005). Developmental deficits in social perception in autism: The role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience, 23*(2–3), 125–141.

Stiles, J., Reilly, J., Paul, B., & Moses, P. (2005). Cognitive development following early brain injury: Evidence for neural adaptation. *Trends in Cognitive Sciences, 9*(3), 136–143.

VISION: TEMPORAL FACTORS

Perception is modified by changes in visual stimuli that take place over time. For example, the apparent brightness of a stimulus may vary as a function of duration, even if its luminance is constant. Further, the appearance of a stimulus is affected by its timing with respect to other stimuli. For example, a flash of light may not appear to be a flash at all if it is embedded in a series of flashes (a phenomenon called *flicker fusion,* described further in the entry). Also described in this entry are the roles of temporal factors in brightness perception and factors in visual physiology and stimulus visibility.

Roles of Temporal Factors in Brightness Perception

The perceived brightness of a stimulus is not only determined by physical intensity but is also affected by duration. That is, a stimulus can be perceived as more intense by increasing either its physical intensity or its duration. Thus, the literature often refers to the time-integrated contrast energy, c, of a stimulus instead of its physical intensity.

One such relationship is conveyed in Bloch's law. Bloch's law (also called the time-intensity reciprocity law) states that a short-duration visual stimulus that is high in physical intensity (I) can appear as bright as a longer-duration stimulus of lower intensity (so long as both stimuli are of a duration, t, that is shorter than a critical duration, τ): for $t \leq \tau$, $c = I \times t$. The critical duration that temporally limits Bloch's law, τ, becomes shorter as the overall intensity of the stimuli increases. Bloch's law presumably operates due to integrative action of the visual system. However, the neural correlates are not currently known. Some reports suggest that a simple decay of neural activity can account for the effect.

In the Broca-Sulzer effect, perceived brightness also depends on stimulus duration. As the duration of a flashed stimulus increases, so does its perceived brightness, but then it decreases. Brücke and Bartley individually reported that the brightness of a flickering stimulus varies as a function of flash duration (an effect now referred to as the Brücke-Bartley effect). Later studies also investigated the role of flicker rate on the perceived brightness of flickering stimuli. Taken together, flicker rate, stimulus duration, interstimulus interval, and flicker on–off ratio (duty cycle) are all temporal factors that impinge upon the perceived brightness of a stimulus.

Flicker Fusion

The temporal factors that affect the appearance of a stimulus include the stimulus duration, but also its temporal interactions with other stimuli. Flicker fusion is the perceived continuous appearance of a flickering light. Although artificial electric lights appear to be stable light sources, in fact their light emission turns on and off cyclically at a rate of 50 to 60 Hz. Movies, computer monitors, TVs, and other presentation devices similarly flicker, despite the fact that they appear continuous. The apparent continuity of these devices is of central interest to the field of temporal visual processing, as well as to understanding how humans interact with the many devices and environments that make up modern society. For flicker fusion to occur, the rate of flicker of the source must be higher than the "critical flicker fusion threshold." This rate is typically 30 Hz or more, but it may vary as a function of context and lighting conditions.

Flicker fusion is thought to occur as a result of a process called "persistence of vision." In 1824, Peter Mark Roget (who also wrote the famous thesaurus) first presented the concept of "persistence of vision" to the Royal Society of London, as the ability of the retina to retain an image of an object for 1/20 to 1/5 a second after its removal from the field of vision. A second principle—the "phi phenomenon" or stroboscopic effect—is closely related to flicker fusion. The effect was first studied between 1912 and 1916 by Max Wertheimer (the founder of Gestalt psychology) and Hugo Munsterberg, who found that subjects perceptually bridge the temporal gap between two consecutive stimulus presentations, allowing them to perceive a series of static images as continuous movement.

Role of Temporal Factors in Visual Physiology and Stimulus Visibility

Temporal factors in visual perception are a function of the response dynamics of neurons in the visual system. Thus, to understand how temporal factors

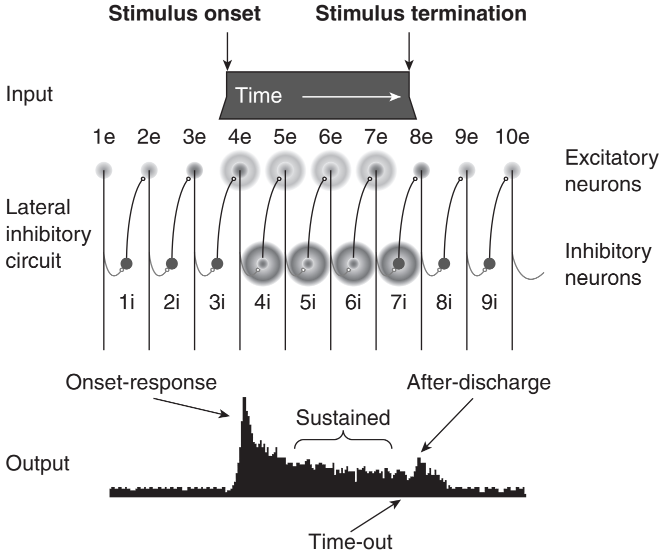


Figure 1 Temporal Response Dynamics and Response to Presentation of a Visual Stimulus

Notes: Temporal response dynamics of one excitatory and one inhibitory visual neuron, followed before, during, and after the presentation of a stimulus in which the stimulus is off (times 1, 2, and 3), on (times 4, 5, 6, and 7), and then off (times 8, 9, and 10). The neuronal response histogram of the excitatory neuron at the bottom shows the average response from 28 neurons in primate area V1 when visually stimulated for 500 milliseconds.

modulate perception, we must determine how visual neurons dynamically respond to stimuli.

Visual cortical neurons follow a stereotypical four-part temporal response dynamic in response to the presentation of stimuli (Figure 1): one excitatory neuron (at times 1e through 10e) and its connected inhibitory neuron (at times 1i through 9i). At times 1e, 2e, and 3e (before the stimulus is presented), there is no excitatory input, so the output remains flat. At time 4e (just after the stimulus, such as a bar of light, is presented), the neuron is excited, causing an onset-response. This leads to the activation of the inhibitory neuron at time 4i, after a slight delay. The inhibitory neuron then feeds back on the excitatory neuron and causes its activity to be suppressed at time 5e. This state of excitatory-inhibitory equilibrium is called the sustained period, which continues through time 7e, after which the stimulus is extinguished. Despite the stimulus having been terminated, the neuron at time 7i is nevertheless activated by the excitatory neuron at time 7e, due to the delayed effect of inhibition. Thus, the excitatory neuron at time 8e is inhibited while not being excited by visual input, and so is in a state of deep suppression called the time-out period, which in turn causes the inhibitory neuron at time 8i to be deeply suppressed due to lack of input. The excitatory neuron at time 9e then exhibits a disinhibitory rebound called the after-discharge due to lack of baseline inhibition (despite the fact that there is no excitatory input).

The Nobel laureate Edgar Adrian and his colleague Rachel Mathews, in the first microelectrode recordings from any visual system (eel optic nerve), found this temporal pattern of activity in response to the presentation of a disk of light. This response dynamic has been replicated in numerous animal models, including primates. Its perceptual consequence is that a visual stimulus appears more salient when it turns on and off than during its midlife.

Flicker fusion is also explained by these dynamics. When two stimuli are presented in close succession, the physiological after-discharge from the first stimulus and the onset-responses from the second stimulus may inhibit each other, thus generating perceptual flicker fusion.

Stephen L. Macknik and Susana Martinez-Conde

See also Contrast Enhancement at Borders; Lateral Inhibition; Visual Illusions; Visual Masking

Further Readings

Bartley, S. H. (1937). The neural determination of critical flicker frequency. *Journal of Experimental Psychology, 21*(6), 678–686.

Bloch, A. M. (1885). Experience sur la vision. *Comptes Rendus de Seances de la Societe de Biologie, Paris, 37,* 493–495.

Macknik, S. L., & Martinez-Conde, S. (2004). The spatial and temporal effects of lateral inhibitory networks and their relevance to the visibility of spatiotemporal edges. *Neurocomputing, 58–60C,* 775–782.

Macknik, S. L., Martinez-Conde, S., & Haglund, M. M. (2000). The role of spatiotemporal edges in visibility and visual masking. *Proceedings of the National Academy of Sciences USA, 97*(13), 7556–7560.

VISUAL ACUITY

Visual acuity is a measure of the keenness of sight. It has been studied, measured, and analyzed for thousands of years because it represents a fundamental limit in our ability to see. Consequently, visual acuity has been used as a criterion for military service and various other occupations, driving, and for receiving social security benefits. Visual acuity is limited primarily by the optics of the eye and by the anatomy and physiology of the visual system. Eye doctors assess visual acuity to assess the optical and physiological state of the eye and visual pathways. This entry concentrates on defining and specifying visual acuity, including minimum visible, resolvable, recognizable, and discriminable acuity.

Defining and Specifying Visual Acuity

How do we define the keenness of sight? Over the centuries, there have been a large number of different ideas about how to define and measure visual acuity, and these can be distilled down to four widely accepted criteria:

1. Minimum visible acuity—detection of a feature.
2. Minimum resolvable acuity—resolution of two features.
3. Minimum recognizable acuity—identification of a feature.
4. Minimum discriminable acuity—discrimination of a change in a feature (e.g., a change in size, position, or orientation).

These different criteria actually represent different limits and are determined by different aspects of the visual pathway.

Minimum Visible Acuity

Minimum visible acuity refers to the smallest object that one can detect. How small would that be? Visual acuity is generally specified in terms of the angular size of the image of the target at the retina. Under ideal conditions, humans can detect a long, dark wire (like the cables of the Golden Gate bridge) against a very bright background (like the sky on a bright sunny day) when they subtend an angle of just 0.00014°. As a comparison, your thumb, when viewed at arm's length, subtends an angle of about 2° on your retina, assuming your thumb is about 2 cm across and your outstretched arm is around 57 cm from your eye. It is widely accepted that the minimum visible acuity is so small because the optics of the eye spread the image of the thin line, so that on the retina it is much wider, and the fuzzy retinal image of the wire casts a shadow that reduces the light on a row of cones to a level that is just detectably less than the light on the row of cones on either side. In other words, although we specify the minimum visible acuity in terms of the angular size of the target at the retina, it is actually limited by our ability to discriminate the intensity of the target relative to its background. Increasing the target size, up to a point, is equivalent to increasing its relative intensity.

Minimum Resolvable Acuity

Minimum resolvable acuity refers to the smallest angular separation between neighboring objects that one can resolve. More than 5,000 years ago, the Egyptians assessed visual acuity by the ability of an observer to resolve double stars. There is currently still debate about how best to define and measure resolution. Today, however, the minimum resolvable acuity is much more likely to be assessed by determining the finest black and white stripes that can be resolved. Under ideal conditions, humans with good vision can resolve black and white stripes when one cycle subtends an angle of approximately 0.017° (1 minute of arc). This minimum resolvable acuity represents one of the fundamental limits of spatial vision: It is the finest high-contrast detail that can be resolved. In foveal vision, the limit is determined primarily by the spacing of photoreceptors in the retina. The visual system "samples" the stripes discretely, through the array of receptors at the back of the retina. If the receptors are spaced such that the whitest and blackest parts of the grating fall on separate cones, we should be able to make out the grating. But if the entire cycle falls on a single cone, we will see nothing but a gray field (or we may experience a phenomenon called aliasing, in which we misperceive the width or orientation of the stripes). Cones in the fovea have a center-to-center separation of about 0.008° (0.5 minutes of arc), which fits nicely with the observed acuity limit of 0.017° (because we need two cones per cycle to be able to perceive it accurately), and each foveal cone has a "private" line to a ganglion cell. Rods and cones in the retinal periphery are less tightly packed together, and many receptors converge on each ganglion cell. As a result, visual acuity is much poorer in the periphery than in the fovea.

Minimum Recognizable Acuity

Minimum recognizable acuity refers to the angular size of smallest feature that one can recognize or identify. Although this method has been used since the 17th century, the approach still used by eye doctors today was introduced more than a century ago by a Dutch eye doctor, Herman Snellen, and his colleagues. Snellen constructed a set of block letters for which the letter as a whole was five times as large as the strokes that formed the letter. Eye doctors don't describe acuity in terms of visual angles and cycles. The last time you visited your eye doctor, she may have made you read letters, decreasing the size of the letters until you made several errors. Then she may have told you that your visual acuity was 20/20 (6/6 if you live in the United Kingdom) if your vision was good, or 20/30 (6/9 in the United Kingdom) if you needed glasses, or possibly 20/10 (6/3 in the United Kingdom) if you could read the smallest letters on the eye chart. This strange method for designating visual acuity was described by Snellen in 1862. Specifically, Snellen defined the visual acuity of a patient whose vision is being tested as follows:

$$\frac{\text{The distance at which the patient can just identify the letters}}{\text{The distance at which a person with “normal” vision can just identify the letters}}.$$

In later adaptations of the Snellen test, the viewer was positioned at a constant distance of 20 feet (or 6 meters in the UK), and the size of the letters, rather than the position of the viewer, was altered. So normal vision came to be defined as 20/20. To relate this back to visual angle, a 20/20 letter is designed to subtend an angle of 5 arc minutes (0.083°) at the eye, and each stroke of a 20/20 letter subtends an angle of 0.017° (our old familiar 1 arc minute). Thus, if you can read a 20/20 letter, you can discern detail that subtends 1 minute of arc. If you have to be at 20 feet to read a letter that someone with normal vision can read at 40 feet, you have 20/40 vision (worse than normal). Although 20/20 is often considered the gold standard, most healthy young adults have an acuity level closer to 20/15.

Minimum Discriminable Acuity

Minimum discriminable acuity refers to the angular size of the smallest *change* in a feature (e.g., a change in size, position, or orientation) that one can discriminate. Perhaps the most studied example of minimum discriminable acuity is our ability to discern a difference in the relative positions of two features. Our visual systems are good at telling where things are. Consider two abutting horizontal lines, one slightly higher than the other. It is easy to discern that, for example, the right line is higher than the left (i.e., discriminate the relative positions of the two lines), even from a long way away. This is an example of a class of visual tasks that have been given the label *hyperacuity* by Gerald Westheimer. These tasks all have in common that they involve judging the relative position of objects, and Westheimer coined the term hyperacuity because, under ideal conditions, humans can make these judgments with a precision that is finer than the size or spacing of foveal cones (a visual angle of approximately 0.008°).

The smallest misalignment that we can reliably discern is known as *vernier acuity*—named after the Frenchman Pierre Vernier, whose scale, developed in the 17th century, was widely used to aid ships' navigators. The success of the vernier scale was based on the fact that humans are adept at judging whether nearby lines are aligned or not. Thus, vernier alignment is still widely

Table 1 Types and Limits of Acuity

Type of Acuity	*Measured*	*Acuity (degrees)*
Minimum visible	Detection of a feature	0.00014
Minimum resolvable	Resolution of two features	0.017
Minimum recognizable	Identification of a feature	0.017
Minimum discriminable	Discrimination of a change in a feature	0.00024

used in precision machines, and even in the dial switches in modern ovens. So how small is it? Under ideal conditions, vernier acuity may be just 0.0008°! This performance is even more remarkable when you consider that it is about 10 times smaller than even the smallest foveal cones. Note that the optics of the eye spread the image of a thin line over a number of retinal cones, and that the eyes are in constant motion, and this performance appears even more remarkable.

As remarkable as the hyperacuities (sometimes also called "position" acuities) might seem, they do not defy the laws of physics. Bill Geisler has calculated that if you placed a machine (known as an ideal discriminator) at the retina, and this machine knew precisely the pattern of photons absorbed by the retinal photoreceptors when the stripes were aligned, and the pattern of photons absorbed when they were misaligned, that this machine could actually perform a lot better than even the best humans. So, the information about the vernier offset is present in the pattern of photons absorbed by the photoreceptors; however, humans must be able to interpret the information despite the constant motion of the eyes. Thus, hyperacuity must ultimately be limited by neurons in the visual cortex that are able to interpolate positional information with high resolution.

Table 1 summarizes the different forms of acuity and their limits.

Dennis M. Levi

See also Perceptual Development: Visual Acuity; Tactile Acuity

Further Readings

Geisler, W. S. (1984). Physical limits of acuity and hyperacuity. *Journal of the Optical Society of America, A., 1*(7), 775–782.

Hecht, S., & Mintz, E. U. (1939). The visibility of single lines at various illuminations and the retinal basis of visual resolution. *Journal of General Physiology, 22*(5), 593–612.

Wade, N. J. (2007). Image, eye and retina (invited review). *Journal of the Optical Society of America, 24*(5), 1229–1249.

Westheimer, G., (Ed.). (1981). Visual acuity. In R. A. Moses, (Ed.), *Adler's physiology of the eye: Clinical application* (7th ed.). St. Louis, MO: C.V. Mosby.

Westheimer, G. (1987). Visual acuity and hyperacuity: Resolution, localization, form. *American Journal of Optometry & Physiological Optics, 64*(8), 567–574.

Westheimer, G. (2005). The resolving power of the eye. *Vision Research, 45*(7), 945–947.

VISUAL CATEGORIZATION: PHYSIOLOGICAL MECHANISMS

Visual categorization refers to the ability to group stimuli according their behavioral significance or meaning. Stimuli from the same category are treated as "similar," though they may differ greatly from one another in their physical features (e.g., a "giraffe" and "lizard" are both of the category "animal"). Likewise, stimuli in different categories may be physically similar, though they are treated as "different" (e.g., a "green apple" and "green tennis ball" are in different categories). Humans and other advanced animals have a remarkable capacity to learn and recognize visual categories, and this ability forms the foundation of complex behavior. Without the ability to learn and recognize the behavioral significance of stimuli, we would be unable to successfully interact with and adapt to the ever-changing world around us. Recent work has shed new light on the brain mechanisms underlying visual categorization and suggests that neurons in the parietal, temporal, and frontal lobes may play a key role in encoding learned information about the category membership of visual stimuli. This entry will explore recognition and categorization, neuronal mechanisms, and future research relating to the physiological mechanisms of visual categorization.

Recognition and Categorization

In our daily interactions with the world, we are faced with a constant stream of incoming sensory stimuli, such as faces, vehicles, furniture, and household objects. In order to plan successful behaviors that move us closer to achieving our goals, the brain is faced with the challenge of making sense of this dizzying array of objects and events around us. Because the brain typically solves this problem with remarkable accuracy and efficiency, allowing us to easily make sense of stimuli, we usually take this ability for granted. However, from a computational point of view, how to recognize the category, or meaning, of stimuli is an extraordinarily difficult problem to solve. In fact, computer scientists working with the most powerful modern computers have made surprisingly little progress in creating systems that can solve even basic recognition tasks, such as recognizing familiar faces or objects. Furthermore, the problem becomes even more difficult when you consider that objects can be viewed from multiple viewpoints, under different illumination conditions, and embedded in cluttered scenes. Thus, the human brain is by far the most sophisticated and successful recognition system in existence. Over the past several decades, neuroscientists have made some progress toward understanding the brain mechanisms underlying recognition and categorization, particularly in the visual system.

Consider the sheer number of unique visual stimuli that an average human observer can recognize, and consider also that each of these stimuli can be viewed from different vantage points, from different distances, and with different illumination. These manipulations produce a massive number of physically unique images for any one object, driving the total number of unique images toward the infinite. How does the brain make sense of this enormous range of images that we might encounter? One strategy might be for the brain to learn and store each possible image separately as "templates," and to attempt to recognize incoming stimuli by searching for an exact match to one of the stored templates. However, the nearly infinite set of stimuli that we encounter would necessitate a nearly infinite storage capacity—something the brain clearly does not have. Furthermore, whereas this strategy might allow us to recognize an image that is exactly identical to one that we had seen in the past, we might be unable to recognize a stimulus if viewed from a slightly different vantage point or under different illumination conditions. Instead, the brain divides the nearly infinite set of possible images into discrete groups or categories.

Of course, we are not born with a large built-in library of thousands of object categories, like "vehicle" and "furniture," that we are preprogrammed to recognize. Instead, we acquire the ability to recognize most categories through learning and experience. A key issue is to understand the mechanisms by which the brain learns and recognizes visual categories. While a true understanding of this process will ultimately require a detailed cellular- and circuit-level description of how such information is learned and encoded, a first step is to identify brain regions that are likely to be centrally involved in the categorization process.

Neuronal Mechanisms

A number of recent studies have provided new insights into the neuronal mechanisms underlying visual categorization. Most of these studies were conducted using the technique of extracellular neuronal recordings from neurons in various brain areas in highly trained rhesus monkeys. This technique has been one of the most commonly used techniques to study questions about visual system function and the relationship between neuronal activity and behavior. This is because the organization of the visual system is similar between monkeys and humans, and monkeys' visual acuity, color vision, eye movements, and discrimination thresholds are similar to that in humans. Furthermore, rhesus monkeys are intelligent and interactive animals that can be readily trained to perform complex behavioral tasks, including visual discrimination, categorization, and memory-based tasks. Thus, they are an excellent model system for investigating the neuronal mechanisms underlying visual learning and recognition.

The visual system of both the monkey and human brain is composed of several dozen distinct brain areas that are highly interconnected with one another and organized hierarchically. Broadly speaking, neurons in brain areas at the earliest stages of the visual system (e.g., the primary visual cortex area V1) have been shown to respond selectively, or encode, relatively simple visual features, such as the orientation of an edge within their receptive fields (RFs). As one moves up the cortical hierarchy, neurons tend to prefer progressively more complex visual features. Near the highest levels of the visual hierarchy, neurons in the inferior temporal cortex show exquisite selectivity for highly complex visual stimuli or features, including faces, geometrical shapes, photographs, and natural objects. Because neurons in the earliest stages of the visual system (e.g., area V1) are primarily involved in basic visual feature processing, learning and experience are unlikely to dramatically alter their encoding properties. Instead, learning-related changes are more likely to occur at higher stages of the hierarchy, perhaps in the highest levels of the visual system (e.g., the inferior temporal cortex) or at processing stages one or more steps beyond the visual hierarchy, such as the prefrontal cortex.

Many of the first neurophysiological studies to address the question of visual categorization were focused on the inferior temporal (IT) cortex, located along the anterior ventral surface of the temporal lobe. Both humans and monkeys with damage to this portion of the IT cortex show marked deficits in visual recognition and perceptual learning. In some cases, these deficits can even be category specific. For example, temporal lobe damage in humans has been shown to cause prosopagnosia, a selective deficit in the ability to recognize faces. Subjects with prosopagnosia can typically describe many of the low-level visual features of the face, but appear deficient in recognizing the face as belonging to a specific familiar individual or in discriminating one face from others. These perceptual and behavioral deficits suggested that the IT cortex might play a role in visual learning, recognition, and categorization. Neurophysiological recordings from neurons in the IT cortex gave further support to this hypothesis, as the activity of individual IT neurons was found to be highly selective for complex visual stimuli such as geometric shapes, familiar objects, and photographs. In some cases, the shape selectivity of IT neurons appeared to correspond to object categories, in that a given IT neuron would respond strongly to stimuli of one category, but more weakly to other categories. The clearest and most widely studied example is face selectivity. Many studies have found clusters of IT neurons that respond much more strongly to faces than to other types of stimuli, raising the possibility that face processing relies on a privileged processing mechanism in the IT cortex because such specialization has not been demonstrated for other classes of visual stimuli in the IT cortex.

Recent studies have focused on comparing the relative roles of multiple brain areas in encoding learned categories by recording from multiple brain areas during task performance. In one study conducted by David Freedman in Earl Miller's lab at MIT in collaboration with Tommy Poggio and Maximilian Riesenhuber, monkeys were trained to perform a shape categorization task in which they had to classify computer-generated “morphed” stimuli as “cats” or “dogs.” The stimuli smoothly varied along a continuum of “catness” and “dogness,” and were separated into two categories by a learned category boundary (at the 50% cat–50% dog midpoint). Stimuli that were close to, but on opposite sides of, the boundary were visually similar to one another. Likewise, stimuli could be visually dissimilar but belong to the same category. After several months of training, the monkeys were experts on performing the task, and neuronal activity was recorded from the IT cortex and the prefrontal cortex (PFC). The recordings revealed that IT and PFC played distinct roles in solving the categorization task. PFC neurons showed a more explicit encoding of the category membership of stimuli. Some PFC neurons responded more strongly to dogs, while others preferred cats. However, in both cases, many PFC neurons showed a clear and explicit encoding of the category membership of stimuli—their activity showed little variability between stimuli in the same category and differed sharply between categories. Recordings from the IT cortex revealed a different pattern of results. Shape selectivity in the IT cortex showed a weaker tendency to encode the learned categories than in PFC. Instead, an IT neuron might respond strongly to only some of the cats or dogs. Thus, IT appeared to be more specific in its shape selectivity than in PFC, showing sharp tuning for smaller subsets of visually similar stimuli, but not more abstract category encoding. This finding suggests that the IT cortex is more involved in visual shape feature processing, whereas more abstract information about the category membership is more strongly encoded in subsequent processing stages, such as in PFC.

Earlier work on visual categorization focused on the IT cortex and, more recently, the PFC work suggests that the parietal cortex may also play a role in encoding learned visual categories. The parietal cortex is generally thought to play a central role in spatial processing, and decades of work have established that areas within the parietal lobe are involved in spatial functions, such as motion processing, directing spatial attention, and guiding eye and reaching movements. However, recent work from several laboratories suggests that the lateral intraparietal (LIP) area may be involved in encoding nonspatial information about the behavioral significance, or category membership, of visual stimuli. In a series of studies conducted by David Freedman and Jamie Fitzgerald in John Assad's laboratory at Harvard Medical School, monkeys were trained to perform a motion-direction categorization task in which visual motion patterns were divided into two categories. Neurophysiological recordings revealed that LIP neurons showed a strong and explicit encoding of the motion categories. However, neurons in the middle temporal (MT) area, an earlier stage of visual motion processing, were strongly direction selective, but did not show a clear and consistent encoding of the learned categories. Recent extensions of that work have suggested that category-like encoding in LIP is also evident for visual shapes as well, suggesting that LIP may play a more general role in encoding the behavioral significance of visual stimuli than had been previously thought.

Future Research

Although recent work has begun to shed light on how the brain learns and recognizes visual categories, important questions remain to be investigated in order to develop a detailed mechanistic understanding. For example, little is known about the brain mechanisms underlying the learning process itself because most prior work (including the previously mentioned studies) examined neuronal activity only after the category–learning process was complete. In addition, because most physiological studies of categorization have focused on visual categorization, almost nothing is known about categorical representations in other sensory modalities and how information is integrated across sensory modalities. Also, it remains to be determined how well laboratory studies of categorization in nonhuman animals relate to real-world categorical perception in more natural viewing conditions and in humans. In addition to studying these open questions with existing techniques, new techniques for monitoring and manipulating neuronal activity—at the level of circuits, cells, and molecules—are needed in order to make meaningful progress toward answering how the brain learns, stores, and recognizes the category membership of visual stimuli.

David J. Freedman

See also Object Perception: Physiology; Visual Processing: Extrastriate Cortex

Further Readings

Bunge, S., & Wallis, J. (2008). *Neuroscience of rule guided behavior.* Oxford, UK: Oxford University Press.

Freedman, D. J., & Assad, J. A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature, 443*(7107), 85–88.

Freedman, D. J., & Miller, E. K. (2008). Neural mechanisms of visual categorization: Insights from neurophysiology. *Neuroscience and Biobehavioral Reviews, 32*(2), 311–329.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2003). A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience, 23*(12), 5235–5246.

Harnad, S. R. (1990). *Categorical perception: The groundwork of cognition.* Cambridge, UK: Cambridge University Press.

Hawkins, J., & Blakeslee, S. (2004). *On intelligence.* New York: Times Books.

VISUAL DEFICITS

See Low Vision

VISION DISORDERS: BLINDNESS

Blindness is the complete absence or severe impoverishment of vision. The World Health Organization (WHO) defines blindness as acuity of less than 20/400 in the better eye (in case of unequal vision loss across the two eyes) or a visual field of less than 10° around fixation. Blindness is to be distinguished from low vision, which is defined as best-corrected acuity of less than 20/60, but better than 20/400. While blindness effectively renders vision ineffective for even the most basic tasks (such as obstacle avoidance), low vision permits some coarse visually guided skills, although it compromises more detail-oriented ones, such as reading and driving.

Over 35 million people across the world are blind, with the highest incidence (over 90%) in developing countries. Each year, 1 million blind people are treated, 6 million people die blind, and 8 million new cases are added. Thus, the net yearly increase of the blind population is 1 million. The economic cost corresponding to the loss of productivity caused by blindness is estimated to be in excess of $40 billion per year. This entry explores some of the primary causes of blindness, some scientific studies of perception in the blind, and frontiers in blindness research.

Primary Causes

Blindness can have a variety of etiologies, ranging from damage to the eye to brain trauma. Next, the causes that account for the bulk of global blindness are described.

Cataracts

A cataract is an opacity of the lens. The likelihood of an opacity increases with age. Although the precise causal factors are still not clearly understood, ultraviolet (UV) exposure, genetic predisposers, smoking, and diabetes mellitus are all believed to heighten the risk. Cataract surgery is one of the great success stories of our time. A short procedure that involves removal of the lens done under local anesthesia (for adults) can almost completely restore clear vision.

Macular Degeneration

The macula is the central region of the retina that includes the fovea and that subserves high-acuity vision. Macular degeneration (MD) is a disorder that compromises the macula and thus severely impairs visual perception. The risk for MD worsens with age and prior familial incidence of the condition. Regrettably, no satisfactory preventive or treatment strategy for MD is currently available.

Diabetic Retinopathy

Overgrowth and/or hemorrhaging of retinal blood vessels associated with diabetes mellitus can lead to retinal damage and eventually blindness. Effective interventions exist, including laser-based cauterization of hemorrhaging vessels.

Glaucoma

Glaucoma is chronically increased intraocular pressure, which can lead to damage of the retinal tissue and, especially, the optic disk. Glaucoma might have genetic underpinnings in some individuals, and might be the outcome of trauma in others. Treatments that reduce intraocular pressure are effective in the early stages of glaucoma. Delay leads to permanent damage to the optic nerve fibers, and hence uncorrectable visual loss.

River Blindness

River blindness, or onchocerciasis, is a common cause of blindness in parts of Africa. The disease is caused by a skin-dwelling nematode whose larvae are transmitted across a population via biting black flies. The flies breed in running streams (hence the prevalence of the disease near rivers). The transmitted larvae migrate across a person's body. In the eyes, these larvae cause a host of problems, including corneal scarring, glaucoma, and cataracts, leading eventually to blindness. Population-based drug-distribution programs have been effective in reducing the incidence of onchocerciasis.

Childhood Blindness

Factors that are implicated in childhood blindness include maternal rubella (infection during the first trimester of pregnancy greatly increases the likelihood of blindness in the child), measles (which causes corneal scarring), vitamin A deficiency (which can lead to corneal drying and scarring), conjunctival infections at birth, and retinal detachment associated with premature births.

Based on an analysis of causal factors, it is believed that nearly three quarters of all blindness in the world is either preventable or treatable. Large-scale programs, such as WHO's Vision 2020, have been launched to reduce and eventually eliminate all cases of avoidable blindness within the first few decades of the 21st century.

Scientific Studies of Perception in the Blind

The study of blindness provides a unique opportunity to investigate diverse aspects of sensory/perceptual processing, as well as issues related to brain plasticity and learning. Here, a few representative studies are briefly described.

Blindness Induced Changes in Nonvisual Perceptual Sensitivities

How accurate is the popular belief that the blind can "hear better" than the sighted? Surprisingly, there have been no definitive answers to this question. Several studies have failed to find a consistent enhancement of auditory sensitivities in the blind relative to the sighted. However, a recent report from Pascal Belin and colleagues at the Montreal Neurological Institute shows that when the performance of early blind individuals (blindness onset within the first two years of life) is examined separately from that of late-blind ones, an advantage in musical note perception becomes evident for the former. Late-blind subjects show no such improvement relative to the normally sighted. This result and others like it highlight the plasticity of perceptual processing mechanisms of the brain and also suggest that the initial years after birth might be especially amenable to such neural resource reorganization.

The Importance of Visual Input During the First Few Months of Life

How do the visual abilities of individuals who have been blind for the first few months of their lives and have then gained sight differ from those of their normally sighted counterparts? Answering this question is important for understanding the processes of early visual learning and time constraints on neural plasticity. Daphne Maurer and her colleagues have systematically studied several children who were born with congenital cataracts and were treated within the first half-year after birth. Interestingly, when these children are tested several years after the surgery, they exhibit subtle but consistent impairments in their ability to extract configural information from images for tasks, such as face recognition. These results suggest that early visual information is important for organizing the perceptual machinery needed to properly analyze subtle spatial relationships in images. Exactly how the early visual input contributes to the development of this spatial perception ability remains unclear.

Language Learning by Blind Children

What role does visual information play in bootstrapping the development of linguistic abilities in children? A strong case can be made for vision as a critical modality for grounding linguistic constructs in real-world entities. The use of prepositions and the learning of nouns, for instance, are expected to be greatly facilitated by visual information. Is this really true? A powerful way of addressing this question is to look at language learning in congenitally blind children, and to examine whether and how early language in the blind differs from that in the normally sighted cohort. Pioneering studies along these lines were undertaken by Barbara Landau and Lila Gleitman. Their work and subsequent investigations have demonstrated an impressively intact facility for language learning in the blind, although the acquisition of some entities that do not have immediate sensory counterparts shows a delay. At the very least, these studies demonstrate that visual information is not a crucial building block for basic language learning. In future work, it will be interesting to explore what role, if any, vision plays in specific and more subtle aspects of linguistic skills.

The Feasibility and Mechanisms of Late Visual Learning

The classical notion of "critical periods" suggests that visual learning is possible primarily during the early stages of development. Just how strict is the critical period for visual learning? Is all visual skill acquisition permanently compromised if the brain is deprived of input from the eyes for the first few years of life? Early animal studies of dark rearing have indicated that early visual deprivation has dramatic consequences on subsequent development. The subjects in these studies exhibited profound and permanent deficits in visual development, as would be predicted by the "critical periods" idea. Does this apply to humans as well? This question has been difficult to answer because cases of sight onset late in life have been rare. However, a new initiative launched by Pawan Sinha of MIT, Project Prakash, is allowing a systematic investigation of this issue. The effort identifies congenitally blind children in India and provides them treatment, thus fulfilling a humanitarian need. But, in doing so, it also affords an opportunity for scientists to examine how much visual skill acquisition is possible after several years of blindness. The results so far suggest that the strict notion of critical periods needs to be refined. The treated children are able to acquire several complex visual skills within months of gaining sight. Longitudinal studies of the skill acquisition process have provided insights into the process by which visual learning proceeds in these children, and perhaps, in normal development. Planned neuroimaging studies will allow for a structure–function analysis by revealing which neural activations are correlated with the onset of different behaviorally observed skills.

Cortical Reorganization in the Blind

What happens to parts of the brain that are deafferented (i.e., cut off from inputs)? This is a question of direct import for our understanding of brain plasticity and functional reorganization. Neuroimaging studies of blind individuals provide an excellent way of addressing this issue. In the normal human brain, over 30% of the cerebral cortex is devoted to processing visual information. What happens to this cortical tissue in the brain of a congenitally blind person? Does it sit silent and unused, or is it recruited by the brain for alternative functions? Is there an age beyond which such recruitment cannot be achieved? These and related questions are being addressed using techniques such as functional magnetic resonance imaging. The visual cortex in blind subjects has been reported to exhibit activations corresponding to tactile stimulation as well as linguistic tasks. For instance, touching Braille letters leads to visual cortical activation in the blind. Interestingly, data suggest that these activations are likely to be functionally significant. Temporary disruption of the primary visual cortex via transcranial magnetic stimulation decreases Braille reading performance of blind subjects but not that of blindfolded sighted ones.

Cross-Talk Between Sensory Modalities

We regularly experience the world via multiple sensory modalities. A ball feels smooth to the touch and looks round. A cube has sharp edges in both modalities. A long-standing question that philosophers and brain scientists alike have grappled with is this: Do the different modalities contribute to an amodal description of the world, or are their representations modality specific? This issue has been called Molyneux's question after a letter Molyneux wrote to the English philosopher John Locke. The query has resisted an answer for over three centuries. An answer will be significant not only from the basic science perspective, but also from the applied viewpoint of designing sensory-substitution devices to help the blind. Will the blind be able to interpret visual information if it is translated into a different sensory modality? We do not yet have a definitive answer, but with ongoing studies of people who gained sight late in life, we do now have a chance of resolving this question.

Frontiers in Blindness Research

Blindness is one of the most debilitating handicaps. It impacts all aspects of a person's life, spanning social interactions, education, and career choices. Blind individuals in the developing world have significantly reduced prospects for health and longevity. Several avenues of research need to be pursued to lessen the global incidence of blindness. Medical research is needed to search for treatments for conditions that currently are incurable. Epidemiological research along with novel service delivery methodologies are needed to identify people in need of treatment and bring them into the fold of medical care. Finally, research into assistive devices is needed to develop aids for the permanently blind. Devices already in the pipeline include "smart canes" that use range sensors and vibrotactile actuators to alert blind users of obstacles in their path, GPS augmented navigation systems, and retinal/cortical implants to compensate for missing retinal tissue or optic nerve damage. Also under development are systems for the partially sighted that can enable people to most effectively utilize the residual vision that they possess. These scientific enterprises carry with them the promise of improving the lives of millions of blind individuals across the world.

Pawan Sinha

See also Low Vision; Molyneux's Question; Visual Disorders: Neuropsychological

Further Readings

Bouvrie, J. V., & Sinha, P. (2007). Visual object concept discovery: Observations in congenitally blind children, and a computational approach. *Neurocomputing, 70*(13–15), 2218–2233.

Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R. J., & Belin, P. (2004). Neuropsychology: Pitch discrimination in the early blind. *Nature, 430*(6697), 309.

Landau, B., & Gleitman, L. R. (1985). The development of speech in blind children. *Language and experience: Evidence from the blind child* (pp. 22–50). Cambridge, MA: Harvard University Press.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. (2001). Early visual experience and face processing. *Nature, 410*(6831), 890.

Mandavilli, A. (2006). Visual neuroscience: Look and learn. *Nature, 441*(7091), 271–272.

Morgan, M. J. (1977). *Molyneux's question: Vision, touch and the philosophy of perception.* Cambridge, UK: Cambridge University Press.

Ostrovsky, Y., Andalman, A., & Sinha, P. (2006). Vision following extended congenital blindness. *Psychological Science, 17*(12), 1009–1014.

Sadato, N., Pascual-Leone, A., Grafman, J., Ibanez, V., Deiber, M.-P., Dold, G., et al. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature, 380*(6574), 526–528.

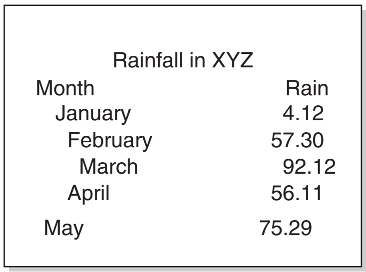
VISUAL DISORDERS: NEUROPSYCHOLOGICAL

See Neuropsychology of Perception

VISUAL DISPLAYS

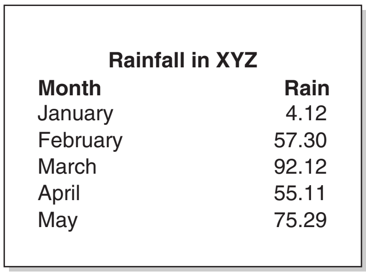
Visual displays are depictions that convey information by means of elements beyond pure text. Examples discussed in this entry include graphic design, drawings, maps, visual interfaces, and data visualization. The common factor in all of these is their reliance on the "visual intelligence" of humans to organize graphically presented information in a way that makes it easier to understand. The design of an effective visual display is based on general principles involving both the nature of human vision and the nature of the task. The particular medium used—paper, canvas, computer monitor—is irrelevant.

The use of visual displays has a long history. Drawings were used tens of thousand of years ago, likely for teaching. With the advent of writing, text became the dominant means of transmitting information and reduced the involvement of visual perception to that of recognizing characters or words. But the use of drawings never



Rainfall in XYZ
Month Rain
January 4.12
February 57.30
March 92.12
April 56.11
May 75.29

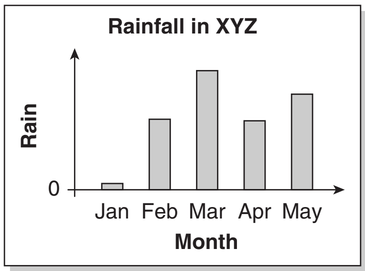
(a) Unorganized Text



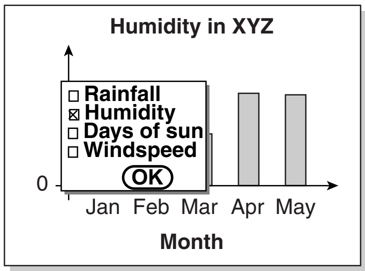
Rainfall in XYZ

Month	**Rain**
January	4.12
February	57.30
March	92.12
April	55.11
May	75.29

(b) Organized Text



(c) Bar Chart



(d) Interactive Chart

Figure 1 Examples of Simple Visual Displays

Notes: Here, the dataset is information about the weather at some location. (a) Unorganized text elements. The display is readable, but is relatively difficult to use. (b) Organized text elements. Organization uses the Gestalt principles of continuity (alignment of text), similarity (common font size and contrast), and proximity. Contrast is used to draw attention first to the primary heading. The result is much easier to read. (c) Bar chart. Data is represented via the height (and area) of the columns. Trends in the data can be easily perceived. (d) Interactive display. Here, the viewer can easily examine different aspects of the information in a data set, allowing him or her to more easily get a sense of the entire situation.

completely disappeared. And displays such as maps were discovered to be a highly effective way of describing the two-dimensional surface of the Earth. More generally, it has been found that when a visual display draws on the appropriate mechanisms of visual perception, it can present information in a way that allows a viewer to understand it far more quickly, accurately, and memorably than if presented by text alone. It has also been found that successful design techniques sometimes point to previously unknown mechanisms.

When used to convey more abstract information by more metaphoric means, visual displays cross over into the domain of fine art. The two domains have a great deal of overlap, drawing on many of the same perceptual mechanisms. They are also compatible—an effective visual display can have great aesthetic appeal. However, the focus of visual display design is on the communication of relatively concrete information, with emphasis on speed and accuracy. As such, different principles are often involved, principles that depend on the nature of the task and on the tradition in which they were developed.

Graphic Design

Graphic design concerns displays centered around text, such as a page or poster. These generally contain graphic elements such as groups, lines, and figures; lines of text may sometimes be quite sparse (Figure 1a). The goal here is to present information to the viewer so they can quickly and easily make sense of the display, with the most important text being perceived first.

An effective graphic design attempts to do this in two ways: (1) send visual attention to the most important points, ideally in appropriate order, and (2) minimize the number of possible candidates by collecting the graphic elements into meaningful groups. Both can be done simultaneously if the display has the appropriate perceptual organization.

Organization can be created by the use of lines or other graphic elements. However, such elements are usually kept to a minimum because they can distract attention. Instead, organization largely involves grouping of the text lines themselves (Figure 1b). Important design principles here include the use of Gestalt laws: alignment of edges (grouping via continuity of the text border), keeping related lines of text close to each other (grouping via proximity), and using the same size and color of font (grouping via similarity). Other principles concern the use of contrast to attract attention, and the use of multiple levels of visual structure; these are often used concurrently. For example, if a title line is at a higher structural level than the supporting text and has a higher contrast, attention will go to it first, allowing the viewer to obtain a summary before exploring further.

Drawings

Drawings convey information about the structure of an object via graphic elements similar to the object they depict. For example, a drawing of a house can have the graphic elements depicting the walls connected to the elements depicting the roof, just as the actual walls are connected to the actual roof. Although text is generally unsuitable for depicting this, it is often used to provide additional information about individual parts. Drawings are believed to engage a form of visual cognition based on mental imagery; a distinction is sometimes made between an "external" drawing in a display and the corresponding "internal" drawing in the viewer's mind.

Several variants of drawings exist. Requiring the graphic elements to be as simple as possible results in *icons,* commonly used in signage and computer interfaces. Removing the need for the graphic elements to resemble real-world objects results in *diagrams,* which allow display of more abstract kinds of information, such as sequences of actions. *Comics* are another variant, in which sequences of drawings are used to convey information about events over time. For all of these, the goal is to present as much understanding of the depicted object as needed. To do this, a drawing must focus visual attention on the important aspects of an item's structure. This is achieved by selection: removing much of the irrelevant detail and highlighting parts of the remainder, for example, by giving particular lines higher contrast. This differs from photographs, where such detail is retained and no highlighting is done. A drawing must ensure that the information is enough to allow the visual system to adequately understand the depicted item.

Understanding a drawing relies heavily on shape perception. Gestalt principles are also useful, both for grouping and figure recognition. The accurate recovery of three-dimensional structure can also be important. This is not always possible from a single view, but a set of drawings (each of a different view) may allow this.

Maps

Maps convey information about the distribution of items or properties—for example, cities or precipitation—over the surface of a world, typically the Earth. With drawings, graphic elements such as lines and regions are dominant, with text only providing information about individual items or parts. Maps also have a strong emphasis on selection: A map generally provides information only about particular aspects of a particular part of a world.

However, maps differ from drawings in several ways. First, except for the occasional need to represent altitude, there is little intent to convey three-dimensional structure. Also, accuracy is often more of a challenge. At large scales, severe distortions arise from projecting the surface of a sphere onto a flat sheet; indeed, it is impossible to accurately portray both direction and area simultaneously. Finally, maps often need to convey information about several dimensions at each location, for example, population density and average income across a city.

Effective maps rely on the use of several perceptual abilities. Shape perception and spatial perception support the understanding of two-dimensional structures, such as rivers and boundaries. Another perceptual aspect is attentional crowding: The extent to which a high density of symbols can be easily understood. Perceptual considerations also influence the choice of graphic variables—for example, which properties should represent cities of different sizes. Effective maps use properties that correspond to basic visual features, such as size, shape, orientation, and hue. These allow for immediate grouping of regions with similar properties and also allow visual attention to easily select particular subsets of properties.

Visual Interfaces

Visual interfaces are displays used for the control of a process. A common example is *human-computer* *interaction,* where the interface is used to control the actions of a computer. Here, the display not only provides the operator with information (e.g., files on the desktop), but also supports interactions with it (e.g., opening folders), as well as control of operations beyond the computer itself (e.g., sending e-mail). More generally, visual interfaces can support effective *human-machine interaction,* in which a human operator controls a machine—perhaps a computer, car, or aircraft.

To be effective, interfaces must make good use of a wide range of perceptual abilities. In addition to the visual intelligence used for other kinds of display, perceptual processes must also support effective action. An example of this is situation awareness, the ability to combine pieces of information into an understanding sufficient to determine the appropriate action. For example, combining information from different views (via mirrors or cameras) determines whether it is possible to turn into an adjacent lane.

Other relevant abilities are the initiation and control of motor movement. To support initiation, the display of the command must be such that engaging it appears to be a natural way to act. The display should also support the use of automatic actions that do not interfere with other operations. If designed well, the resulting interface will be *transparent,* allowing the operator to focus on the task itself rather than the tools. An interesting possibility here is to design interfaces for the "how" stream in vision, a nonconscious system concerned with the real-time control of actions, which may be largely separate from the system underlying our conscious experience of the world.

Data Visualization

Data visualization displays present data sets in a form intended to allow the viewer to rapidly perceive any patterns they may contain, such as trends and outliers. Examples are graphs and scatter plots. The goal here is to "use vision to think"—to use the visual intelligence of the viewer to analyze the contents of the display, reducing the need to carry this out at higher cognitive levels.

Several forms of data visualization have been developed. Graphs and bar charts are perhaps the simplest of these, based on static graphic elements and typically displaying a medium-sized data set of tens or hundreds of items. In a graph, each item in the set is represented by an element with a location corresponding to the information it contains—for example, the number of sales for a particular month. A bar chart is similar, except that the height of the bar replaces vertical position (Figure 1c). Multiple properties per item can be represented by multiple graphic properties, such as the size and color of the elements. Relevant perceptual abilities include attentional selection of particular categories, grouping of similar items into clusters, and the perception of the overall shape and orientation of these clusters, which reflect things such as trends and correlations. Interactive displays have also been developed, allowing different aspects of a data set to be easily examined (Figure 1d). Perceptual abilities having to do with motor control and navigation in virtual spaces are relevant here.

Specialized displays have recently been developed for large data sets. In *scientific visualization,* data is displayed as a dense image, with each point containing information about physical properties that are not necessarily visible in their own right—for example, the density, pressure, and temperature of a fluid within an engine. Successful interpretation uses several kinds of visual intelligence, including edge detection, color perception, and shape perception. Meanwhile, *information visualization* is concerned with the display of more abstract data, such as communication networks or financial predictions. Many of the same principles and perceptual mechanisms are also relevant here. Because consideration is no longer restricted to the physical world, nonintuitive patterns are more often encountered, requiring the mapping to perceptual abilities to be pushed to its limits.

Another recent development is the emergence of *visual analytics.* Displays here not only allow the visual system to use its own intelligence, but also support higher level thinking when the lower level mechanisms are insufficient. For example, allowing an analyst to explore the incidence of a disease as a function of various factors (e.g., geographical location, social network) might help determine important transmission factors, even if the relevant patterns are not immediately evident from the visual display alone. More generally, visual analytics systems are intended to provide a seamless coupling of the various kinds of human intelligence with various kinds of machine intelligence, enabling a single user to make sense of immense amounts of imperfect, complex data.

Ronald A. Rensink

See also Attention: Spatial; Computer Graphics and Perception; Gestalt Approach; Perceptual Organization: Vision; Pictorial Depiction and Perception; Vision: Cognitive Influences; Visual Search

Further Readings

MacEachren, A. M. (1995). *How maps work: Representation, visualization, and design.* New York: Guilford Press.

Massironi, M. (2002). *The psychology of graphic images: Seeing, drawing, communicating.* Mawah, NJ: Erlbaum.

McCloud, S. (1993). *Understanding comics: The invisible art.* New York: Paradox Press.

Monmonier, M. (1996). *How to lie with maps* (2nd ed.). Chicago: University of Chicago Press.

Raskin, J. (2000). *The humane interface: New directions for designing interactive systems.* Reading, MA: Addison-Wesley.

Rensink, R. A. (2007). The modeling and control of visual perception. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 132–148). New York: Oxford University Press.

Thomas, J. J., & Cook, K. A. (2005). *Illuminating the path: The research and development agenda for visual analytics.* Richland, WA: National Visualization and Analytics Center.

Tufte, E. R. (1992). *The visual display of quantitative data.* Cheshire, CT: Graphics Press.

Ware, C. (2004). *Information visualization: Perception for design* (2nd ed.). New York: Morgan Kaufman.

Williams, R. (2004). *The non-designers design book* (2nd ed.). Berkeley, CA: Peachpit Press.

VISUAL FILLING IN AND COMPLETION

Object perception crucially depends on accurate perception of *object boundaries and surfaces.* Boundary perception depends on *boundary completion processes.* Surface perception depends on *surface filling-in processes.* Why filling-in and completion are necessary and how they work are the topics of this entry. To simplify, it is assumed that surfaces and boundaries are defined only by differences in brightness.

Why Filling In and Completion Are Necessary

The need for boundary completion can be understood when considering the following example. If one stands under a tree and looks up while holding a pencil in front of one's eyes, the tree's branches and pencil will merge into a single complex distribution of dark and light on the retina. On the retina, the branches are "attached" to the pencil, but clearly, this is not what is perceived. Somehow the visual system "knows" that the branches of the tree continue behind the pencil in front, despite the fact that the continuation of the tree branches is not present in the retinal input. How the brain determines that the two parts of a tree branch occluded by a pencil (or another object in front) belong together is the problem of boundary perception.

How Filling In and Completion Work

Introducing a few basic concepts by means of metaphor will fully elucidate the problem of boundary perception, and will also give insight into the problem of surface perception. The visual system can be compared to a camera (the "retina") in which tiny light sensors each send an electrical signal to a pixel on a monitor (e.g., area V1, the primary visual cortex). This metaphor is based on the finding that there is a roughly point-by-point anatomical connection between locations on the retina and sites of electrical activity in V1 (V1 is therefore called a retinotopic map). However, to complicate matters, there is only a small minority of neurons in V1 that simply represent light values

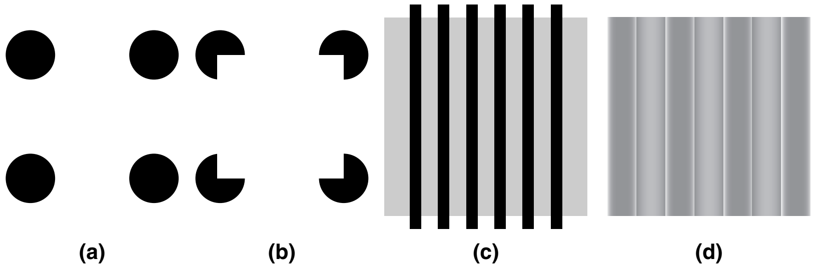


Figure 1 Illustrations of Boundary Completion and Surface Filling-In

Sources: Investigators who introduced these figures to experimental research are Kanizsa (Figures 1a and b), and Craik, O'Brien, and Cornsweet (Figures 1c and d).

(or brightness) in specific retinal locations. Instead, most neurons in V1 represent local contrast in the image. Using the monitor metaphor, most pixels in the monitor are connected to two neighboring light sensors in the camera, rather than just one. These pixels are only turned on if two neighboring light sensors in the camera signal different light values, and are not turned on when the neighboring light sensors signal the same light value. This circuitry in the visual system emphasizes contrasts (often related to boundaries), while most information about surfaces is lost; hence the additional problem of surface perception. The primary role of boundaries in object recognition is revealed by the ease with which we interpret line drawings and cartoons, in which recognition is based only on boundaries and not on the filling in of surfaces.

The problem of boundary and surface perception can now be summarized by asking the questions how do individual neurons in the visual system cooperate to link local information that belongs to the same boundary, and how do they cooperate to reconstruct the percept of surfaces that seems to be lost as a price for the ability to see local contrasts. In essence, the solution to these problems resides in an exchange of information between neighboring neurons (through anatomical connectivity that goes beyond the scope of this entry).

Based on empirical evidence and theoretical models, three types of information exchange between neurons have been proposed. First, neurons activated by local contrasts form a network (the boundary completion system) within which information is exchanged that permits the completion of gaps in object boundaries caused by the presence of objects in front. Second, in the sparser network of neurons that can represent light values (the surface filling-in system), brightness information close to the boundary is interpolated toward the middle of each surface (surface filling-in). Hence, contrast at the border is considered more important for the perception of surfaces than the physical reflection of light from the middle of the surface (see Figures 1c and d). Third, neuronal boundary representations in the boundary completion system keep the spread of surface feature (e.g., brightness) in the surface filling-in system within the confines of the boundaries as represented in the boundary completion system. Returning to the monitor metaphor, contrast-related pixels complete information along aligned fragments of contours, light-related pixels spread brightness, and this spread is stopped in appropriate places by a signal from contrast-related pixels. A large body of evidence supports the idea of separate but interacting systems for boundary completion and surface filling-in, extending from V1 into extrastriate areas.

In Figures 1(a) and (b), the presence of boundary interpolation is illustrated with the Kanizsa square. Figure 1(a) shows four disks that in Figure 1(b) appear to be overlaid by a white square. Interestingly, the suggestion of a square and its contours is derived entirely from the alignment of edges in Figure 1(b). The visual system constructs the boundaries of a square bridging the gaps between edge segments, although neither the square nor its boundaries are physically present. This illusion is considered support for a mechanism that automatically completes boundaries based on local information. Note that this automatic process is based on "statistical knowledge" about the outside world. Specifically, the visual system assumes that a precise alignment of edges in the retinal light distribution is unlikely, except if they are related to a single, real contour in the outside world. Hence, gaps between aligned edges are completed, and a square on top of 4 disks is assumed. Note that the completion of the boundaries of the Kanizsa square is accompanied by a surface filling-in process that leads to the square being perceived as brighter than the surrounding white background. In a related process, the black disks are completed "behind" the reconstructed surface. Filling-in and completion processes related to visible figures (the Kanizsa square) are referred to as "modal," while these processes are referred to as amodal when related to invisible parts of figures (the parts of the disks occluded by the Kanizsa square assumed in front by the visual system).

Figures 1(c) and (d) illustrate the existence of surface filling-in processes. Figure 1(c) shows that the physical gray levels in between the black bars are the same. When the black bars are removed, some regions in the stimulus are seen as much brighter and others as darker, an illusion referred to as the Craik-O'Brien-Cornsweet illusion. The difference in brightness in Figure 1(d) is due to the luminance difference (contrast) at the edges of the surfaces. Surfaces outlined by bright edges are perceived as much lighter than surfaces outlined by dark edges. This supports the idea that the surface features that define the edges are interpolated to reconstruct the percept of the surface (in this case brightness). Boundary and surface completion processes are also hypothesized to underlie filling in across the blind spot caused by the absence of retinal receptors where the optic nerve leaves the eye.

Future Research

The present entry on the topic of completion and filling-in cannot adequately reflect the full complexity of this research field. To do justice to some of the ongoing discussion on this topic, it must be mentioned that while the mechanisms of boundary completion are generally agreed upon, the view on surface filling-in in the present entry represents dominant concepts that are not shared by all investigators. For example, it has been suggested that a spread of surface information from the borders to its middle in retinotopic maps is not required. Instead, a mechanism might exist that simply "assumes" that surface properties just inside its edges should also be present in its middle. This is a symbolic (logical) operation that might be carried out in the high-level visual cortex. Some studies have failed to find evidence for the spread of information thought to underlie surface perception in early retinotopic visual areas, and this has been taken by some as support for symbolic theories of surface perception. Hence, further empirical work is required to test the traditional view as well as symbolic theories, and to resolve some inconsistencies in the findings among some studies. Because the traditional view and symbolic theories are not mutually exclusive, elements from both theories might find support in the future.

P. De Weerd

See also Consciousness; Gestalt Approach; Object Perception; Object Perception: Physiology; Perceptual Organization: Vision

Further Readings

Denett, D. (1991). *Consciousness explained.* Boston: Little, Brown.

De Weerd, P. (2006). Perceptual filling-in: More than the eye can see. *Progress in Brain Research, 154*(1), 227–245.

Komatsu, H. (2006). The neural mechanisms of filling-in. *Nature Reviews Neuroscience, 7*(3), 220–231.

Pessoa, L., & De Weerd, P. (Eds.). (2003). *Filling-in: From perceptual completion to skill learning.* Oxford, UK: Oxford University Press.

VISUAL ILLUSIONS

Visual illusions are subjective percepts that do not match the physical reality of the world. When we experience a visual illusion, we may see something that is not there, fail to see something that is there, or see something different from what is there. Visual illusions not only demonstrate the ways in which the brain fails to recreate the physical world, but they are also useful tools to identify the neural circuits and computations by which the brain constructs our visual experience.

The terms *visual illusion* and *optical illusion* are often used interchangeably. However, unlike visual illusions, optical illusions do not result from brain processes. Instead, an optical illusion is the perception of a distortion that results from the physical properties of light, such as reflection and refraction, and/or the optics of the eye. An example of an optical illusion is the phenomenon in which a pencil looks bent when it is placed upright in a glass of water, owing to the differing refraction indices of air and water. An example of a classical visual illusion is the Ebbinghaus illusion, named for its creator, Hermann Ebbinghaus. If two identical circles are placed side by side, one surrounded by large circles and the other surrounded by small circles, the first central circle will look smaller than the second one (see *Action and Vision,* Figure 2a). The Ebbinghaus illusion cannot be explained by the physical properties of the visual stimulus or by the optics of the eye. Instead, it is due to neural processes that compare a visual object with its context.

Only a fraction of the visual illusions known today have been developed within the framework of the visual sciences. Visual artists have often used their insights regarding perception to create visual illusions in their artwork. Historically, long before visual science existed as a formal discipline, artists had devised a series of techniques to "trick" the brain into thinking that a flat canvas was three-dimensional or that a series of brushstrokes was in fact a still life. Thus, the visual arts have sometimes preceded the visual sciences in the discovery of fundamental vision principles. In this sense, art, illusions, and visual science have always been implicitly linked. This entry describes various types of visual illusions.

How to Make Visual Illusions

Some visual illusions are developed intentionally by applying known visual principles to stimuli patterns and/or experimenting with variations of existing illusions. Other illusions are discovered completely by chance: An attentive observer may simply notice something strange about the way that the world looks and try to understand and replicate the underlying conditions leading to the unusual percept. Finally, illusions may be discovered through the application of known physiological principles of visual processing in the brain.

One example of this last method is the standing wave of invisibility, a type of visual masking illusion in which the visibility of a central bar (the target) is decreased by the presentation of flanking bars (the masks) that flicker in alternation with the target. This illusion was predicted (by Stephen Macknik and Margaret Livingstone) from the responses of visual neurons to flashing objects of varying durations. The Standing Wave of Invisibility illusion demonstrates that a set of masks can render a target perpetually invisible, even though the masks do not overlap the target spatially or temporally. The invisibility of the target results from the adjacent masks suppressing the neural responses normally evoked by the onset and the termination of the target.

Categories of Visual Illusions

Some attempts have been made to classify visual illusions into general categories with varying degrees of success. One substantial obstacle to classifications or taxonomies of visual illusions is that some visual illusions that seem similar may be due to disparate neural processes, whereas other visual illusions that are phenomenologically different may be related at a neural level. Taking these shortcomings into account, some representative categories and examples of visual illusions follow. When known, their underlying neural bases are also discussed. However, the reader should keep in mind that the neural underpinnings of many visual illusions—especially those discovered recently—are not understood. What follows is by no means an exhaustive list.

Adaptation Illusions

The first documented visual illusion was described in Aristotle's *Parva Naturalia.* This illusion, later known as the "waterfall illusion," can be observed while looking at a waterfall, river, or other flowing water. Watch the flowing water for a while (a minute or more works best), and then quickly shift your center of gaze to the stationary objects next to the water (for instance, the rocks to the side of the waterfall). The stationary objects will appear to flow in the opposite direction to that of the water. The illusion occurs because neurons that detect motion in a specific direction (for instance, downward motion if you stare at a waterfall) become adapted (that is, less active) in response to steady stimulation. Neurons that have not been adapted (such as the neurons that detect upward motion) are more active in comparison, despite having been at rest. The differential responses of both neuronal populations produce the illusion of the stationary rocks to the side of the waterfall flowing upward for a few seconds.

Brightness Illusions

Some visual illusions change the apparent brightness of objects. Brightness and color illusions often occur because the brain does not directly perceive the actual wavelength and light reflected from objects in the world. Instead, it compares them to those of other objects in the vicinity. For instance, the same gray square will look lighter when surrounded by black than when surrounded by white. Thus, for the brain, perception is often context dependent.

The Hermann grid is another classic example of a brightness illusion. In this phenomenon, a white grid against a black background shows dark illusory smudges in the intersections. Conversely, a black grid against a white background results in whitish smudges perceived at the intersections (see *Contrast Perception,* Figure 3a). In 1960, Günter Baumgartner measured the responses of visual neurons during the presentation of Hermann grid stimuli. He concluded that the illusion is due to differences in the firing of center-surround retinal ganglion cells to the various parts of the grid (intersecting versus nonintersecting regions). Thus, the Hermann grid illusion has been traditionally interpreted as a perceptual result of lateral inhibition. However, recent research suggests that the retinal ganglion cell theory is incomplete and that the illusion may be generated at the cortical level.

Color Illusions

These are illusions that modify the apparent color of an object. Some classical color illusions are based on simultaneous color contrast. For instance, a gray circle will take on a reddish hue when placed against a green background, and a greenish hue when placed against a red background. This local contrast effect is based on retinal lateral inhibitory processes. Other context-dependent color illusions, such as the “Rubik cube” created by R. Beau Lotto and Dale Purves, are more difficult to explain by local lateral inhibition at the level of the retina, and may thus reflect a more central origin.

Benham’s disk, or Benham’s top, was discovered in 1894 by C. E. Benham, a toymaker. A spinning top with a certain pattern of black and white lines appears to take on colors as it rotates. This illusion has been studied by vision scientists for over 100 years, and it continues to inspire novel research. The underlying neural processes are not well understood, but current theories point toward retinal circuits.

Illusions of Size

The apparent size of an object is changed, usually due to contextual cues. In the Ponzo illusion, two horizontal lines of the same length are superimposed on a pair of converging lines resembling train tracks. The upper line (closer to the converging end of the tracks) seems longer than the lower line (closer to the diverging end of the tracks). The illusion is probably due to the fact that the brain interprets the upper line as farther away than the lower line. The Moon illusion (the perception that the moon looks bigger when close to the horizon than when high up in the sky) might be at least partially related to the Ponzo illusion. That is, the moon close to the horizon may look larger because of accompanying contextual cues, such as trees and houses, indicating that the moon must be far away. Such contextual cues are absent when the moon is high up in the sky.

The Ebbinghaus illusion, discussed in the introduction to this entry, is another example of a classic size illusion.

Shape and Orientation Illusions

These are illusions in which an object appears to take on shapes or orientations that are different from the actual physical ones. Distortion effects are often produced by the interaction between the

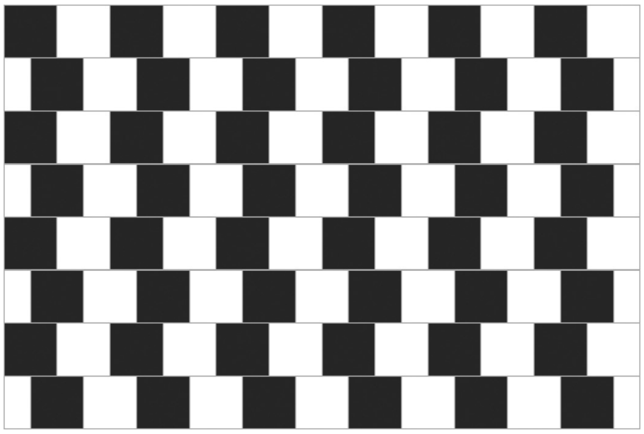


Figure 1 The Café Wall Illusion

actual shape or orientation of the object, and the shapes or orientations of other nearby figures. A classical example is the Café Wall illusion, first discovered in a café in Bristol, England. The black and white tiles in the Café Wall are perfectly straight, but look tilted (see Figure 1).

Invisibility Illusions

In an invisibility illusion, observers fail to perceive an extant object in the physical world. In motion-induced blindness, the observer fixates the center of a display consisting of several stationary circles and a surrounding cloud of moving dots. Although the stationary circles remain physically extant on the display, they fluctuate in and out of visual awareness for the duration of the viewing (sometimes only one circle disappears, sometimes two, sometimes all of them). The neural mechanisms underlying this phenomenon are currently unknown.

The standing wave of invisibility, described earlier, is another example of an invisibility illusion.

Illusory Motion

Some stationary and repetitive patterns generate the illusory perception of motion. The illusory effect is usually stronger if you move your eyes around the figure. If you keep your eyes still, the illusion tends to diminish or even disappear completely. For instance, in the Rotating Snakes illusion created by Akiyoshi Kitaoka, the "snakes" appear to rotate. But nothing is really moving, other than your eyes. If you hold your gaze steady on one of the black dots on the center of each "snake," the motion will slow down or even stop (see Figure 2). Bevil Conway and colleagues showed that the critical feature for inducing the illusory motion in this configuration is the luminance relationship of the static elements. Illusory motion is seen from black to dark gray to white to light gray to black. When presented alone, all four pairs of adjacent elements each produced illusory motion consistent with the original illusion. Also, direction-selective neurons in macaque visual cortex gave directional responses to the same static element pairs, in a direction consistent with the illusory motion. These results demonstrated directional responses by single neurons to static displays and suggested that low-level, first-order motion detectors interpret contrast-dependent differences in response timing as motion.

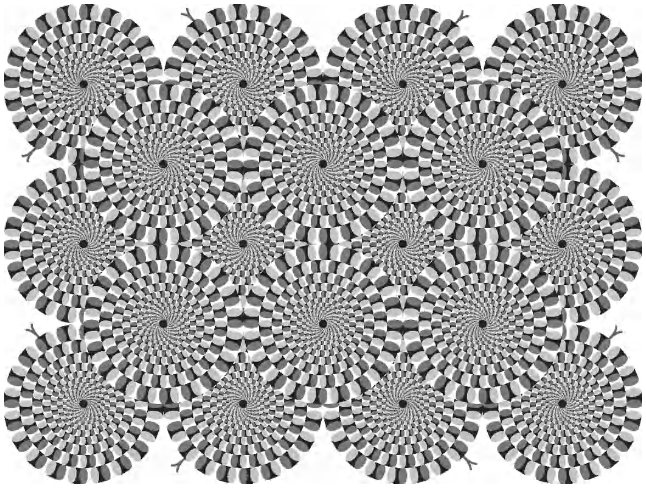


Figure 2 The Rotating Snakes Illusion

Source: Courtesy of Akiyoshi Kitaoka.

Stereo-Depth Illusions

Your left eye and your right eye convey slightly different views of the world to your brain. Close your left and right eye in rapid alternation. You will see that the image shifts left to right. Your brain integrates these two images into a single stereo image, which conveys a sense of depth. This is the principle behind stereo-depth illusions. The wallpaper illusion is a classic example, which arises when observing a pattern of horizontal repetitive elements, such as in wallpaper. If viewed with the appropriate vergence, the repetitive elements appear to float in front or behind the background. The wallpaper illusion is related to the illusions portrayed in the famous Magic Eye books (the Magic Eye illusions are based on a special type of repetitive pattern, called a random dot autostereogram).

Susana Martinez-Conde and Stephen L. Macknik

See also Afterimages; Contrast Enhancement at Borders; Hallucinations and Altered Perceptions; Impossible Figures; McCollough Effect; Nonveridical Perception; Pictorial Depiction and Perception; Visual Masking

Further Readings

Conway, B. R., Kitaoka, A., Yazdanbakhsh, A., Pack, C. C., & Livingstone, M.S. (2005). Neural basis for a powerful static motion illusion. *Journal of Neuroscience, 25*(23), 5651–5656.

Gregory R. L. (2005). The Medawar lecture 2001 knowledge for vision: Vision for knowledge. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 360*(1458), 1231–1251.

Livingstone, M. S. (2008). *Vision and art: The biology of seeing.* New York: Abrams.

Macknik, S. L., & Livingstone, M. S. (1998). Neural correlates of visibility and invisibility in the primate visual system. *Nature Neuroscience, 1*(2), 144–149.

Macknik, S. L., & Martinez-Conde, S. (2008, October). A perspective on 3-D visual illusions. *Scientific American Mind, 19,* 20–23.

Purves, D., Wojtach, W. T., & Howe, C. (2008). Visual illusions: An empirical explanation. *Scholarpedia, 3*(6): 3706.

Troncoso, X. G., Macknik, S. L., Otero-Millan, J., & Martinez-Conde, S. (2008). Microsaccades drive illusory motion in the *Enigma* illusion. *Proceedings of the National Academy of Sciences USA, 105*(41), 16033–16038.

VISUAL IMAGERY

How many windows are there in your home? Is the green of grass darker than the green of pine trees? How many straight and/or curved lines make up an uppercase letter E? If you have not thought about these questions before, it is likely that you experienced a visual image while finding their answers. For example, to determine the number of windows in your home, you might have imagined yourself standing in each room and counting the number of windows you “saw” in your mental image. *Visual imagery* refers to the experience of seeing something that is not physically present, so that there is no corresponding sensory input to your visual system. It is often referred to as “seeing with the mind’s eye.” Most people report that they experience visual images when answering the types of questions posed above, as well as when figuring out how to best pack suitcases in the trunk of their car or rearrange the furniture in their living room. Distinguished scientists and inventors, such as Albert Einstein, Nikola Tesla, and Richard Feynman, reported that their thought processes were accompanied by the experience of mental imagery. For example, Tesla reported that when he first designed a device, he would run it in his head for a few weeks to see which parts were most subject to wear. In what ways is visual imagery like seeing? In what ways is thinking with images different from other forms of thinking? These questions will be discussed in this entry, along with the perceptual characteristics of images, the physiological basis of imagery, the imagery debate, visual versus spatial images, and the functions of visual images.

Perceptual Characteristics of Images

Objective measures have shown that the experience of having a mental image is similar to the experience of seeing in many respects. The time to answer questions about objects in a mental image is related to the relative size of those objects, as if one has to “zoom” into the image to see the details of the object’s appearance. For example, it takes longer to “see” whether a rabbit has whiskers if you imagine a rabbit next to a fly than if you imagine a rabbit next to an elephant. Time to scan between objects in a mental image is also related to the distance between these objects, just as it takes more time to scan between objects that are farther apart when looking at a real scene. For example, if you imagine a map of the United States, it takes longer to scan from San Francisco to New York

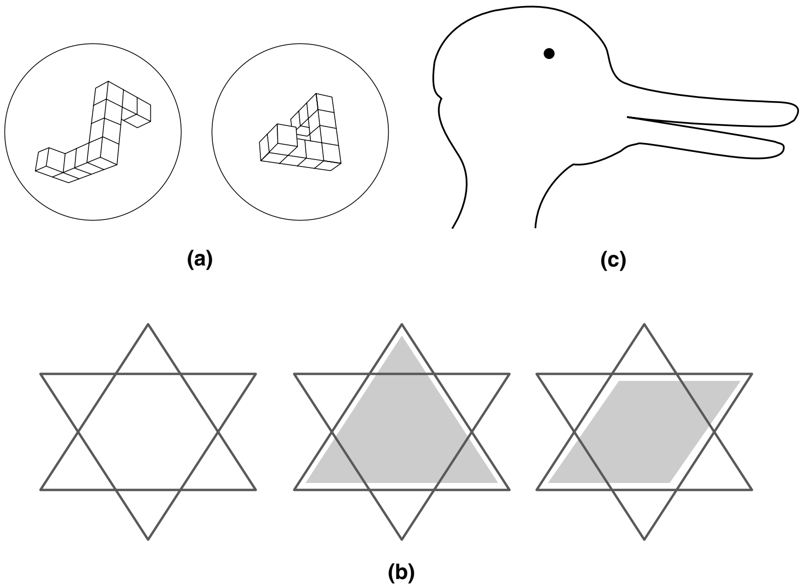


Figure 1 Stimulus Examples

Sources: (a) Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701–703. (b) Reed, S. K., & Johnson, J. A. (1975). Detection of parts in patterns and images. *Memory & Cognition, 3*(5), 569–575. (c) Chambers, D., & Reisberg, D. (1985). Can mental images be ambiguous? *Journal of Experimental Psychology: Human Perception and Performance, 11*(3), 317–328.

than from San Francisco to Seattle. Imagined movements of objects appear to be analogs of the movements themselves. For example, when people are asked to judge whether the two objects in Figure 1(a) depict the same object or are mirror images of each other, time to answer is linearly related to the angular difference in orientation of the two objects, as if people "mentally rotate" the two images into congruence. The larger the angle of rotation, the longer it takes.

Despite these similarities, there are also important differences between mental images and visual percepts. Images are generally less vivid and detailed than percepts. They are also internally organized so that it is easier to "see" certain subcomponents of images than others. For example, when people visualize a Star of David, it is easier for them to see a triangle than a parallelogram in their image, although both shapes are evident in the pictures in Figure 1(b). Furthermore, people cannot always reinterpret patterns in images as well as they can reinterpret them in pictures. For example, when asked to form an image of the picture in Figure 1(c) that is initially presented as a picture of a duck, most people are unable to "see" the alternative interpretation (a rabbit) in their image. However the same people are subsequently able to reinterpret a picture that they draw of their image. This is not to say that images can never be reinterpreted. Imagine a capital letter D, rotate it 90° counterclockwise, and attach a capital J to it from below. What do you see? Most people have no trouble reinterpreting this as an image of an umbrella.

Physiological Basis of Imagery

There is now considerable evidence that visual imagery shares brain mechanisms with visual perception. Brain imaging studies using techniques such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) show that the primary visual cortex (also known as V1 or area 17) and the secondary visual cortex (also known as V2 or area 18), the first areas in the brain to receive input from the retina, are activated during many imagery tasks. The patterns of activation in these areas during imagery have also been shown to mimic perception in at least some ways. For example, when people view objects that subtend smaller visual angles, activation is more posterior in the primary visual cortex than for larger percepts. The same is true of smaller versus larger images. Individuals with brain lesions that affect visual perception are also impaired in visual imagery tasks. People with brain lesions that affect their color vision are also unable to answer questions that depend on imagining colors, such as judging which is darker, the green of grass or the green of pine trees. When repetitive transcranial magnetic stimulation (rTMS) is used to temporarily disrupt the functioning of the visual cortex, people are impaired in tasks that involve comparing visual patterns from memory, whereas this does not occur when rTMS is aimed at nonvisual areas of the cortex. However, not all imagery tasks activate these early visual areas (V1 and V2).

A recent review of neuroimaging studies of mental imagery revealed three characteristics of studies in which activation was found in these areas: (1) the task involved inspecting high-resolution details of images, (2) the task required the visualization of shapes, and (3) the measurement technique was particularly sensitive. Early visual areas are not activated during more spatial imagery tasks, such as mental rotation or imagining taking a walk through your neighborhood.

The Imagery Debate

While the phenomenology of visual imagery is not at issue, there has been much debate in psychology about what is special about mental imagery, how it arises, and the nature of the internal representations or mental codes that underlie the experience of mental imagery. According to the most developed theory of mental imagery, proposed by Stephen Kosslyn and his colleagues, mental images reflect a distinct type of internal representation, a *depictive* representation. A depictive representation of an object resembles that object, such that parts of the representation reflect parts of the object, the shapes of these parts correspond to the shapes of the object, and the spatial relationships between parts of the representation correspond to spatial relationships among the components of the real object. Depictive representations can be distinguished from more abstract propositional representations, which are representations of basic facts. For example, a picture of a cat on a sofa is a depictive representation, but the statements "the cat is on the sofa," "the sofa is soft," "the cat is licking its paws" express propositions.

Kosslyn and his colleagues argue that the activation of the primary and secondary visual cortex during visual imagery tasks is direct evidence for depictive representations in imagery tasks. These areas are topographically mapped, such that specific neurons in these areas represent specific locations in space. That is, space on the cortex represents space in the world. According to Kosslyn's model of mental imagery, object properties are stored in abstract propositional codes in long-term memory. During imagery tasks, these representations are used to recreate depictive representations in a "visual buffer," which is implemented by the set of topographically organized areas of the visual cortex. This image generation process can be thought of as running perception in reverse (given that perception starts with activation in the early visual areas and ends with more abstract representations of object properties). Once an image has been generated, the same perceptual mechanisms that are active during perception can then be applied to inspect the images in the visual buffer and make judgments about their shape, color, and so on.

An alternative account of mental imagery, most strongly argued by Zenon Pylyshyn, proposes that what distinguishes images is their content, not their format. According to this view, there is no reason to postulate a separate form of representation underlying imagery. Visual images are representations of the visual appearance of objects, but thinking with images may depend on the same type of mental code as other forms of thinking, such as thinking with words. Pylyshyn argues that it is important to distinguish between the content and the format of an internal representation, and our subjective experience does not necessarily give us access to the format of our thoughts. Moreover, he suggests that current knowledge of neuroscience does not allow us to make definitive arguments about the format of representations on the basis of brain structures. According to Pylyshyn, when we experience mental imagery, we are simulating what something would look like, based on *tacit knowledge,* or what we have come to know about the appearances of objects from our visual experiences, but our conscious experience of an image may arise from representations that are not themselves depictive.

Visual Versus Spatial Images

It is important to keep in mind that not all imagery tasks have the same demands, so one theory of imagery may not account for all tasks. Some imagery tasks, such as answering questions about the shapes of letters or objects, require people to inspect details of the appearance of objects and are accompanied by activation of the primary visual cortex. Other imagery tasks, such as mental rotation or making judgments about the spatial locations of objects activate parietal rather than the visual cortex and can be performed by congenitally blind as well as by sighted people, suggesting that they may rely on multimodal or amodal representations rather than specifically "visual" representations. Individual difference studies also suggest that there are dissociations between the ability to perform *visual* and *spatial* tasks. For example, there are large individual differences in vividness of visual imagery, as was first studied by Sir Francis Galton when he asked people to imagine their breakfast table. There are also large individual differences in ability to imagine spatial transformations of images, as measured by spatial ability tests. But these individual differences dimensions are unrelated. The dissociation between visual and spatial aspects of mental images is thought to reflect a division of labor in the visual system between an *object* properties pathway (otherwise known as the "what" system) that projects from the occipital to the inferior temporal lobe, and a *spatial* properties pathway (otherwise known as the "where" or "how" subsystem) that projects from the occipital lobe to the posterior parietal lobe.

Functions of Visual Images

Although the imagery debate continues to stimulate important empirical research using both behavioral and neuroscience methods, there are many unanswered questions about mental imagery that depend less on its format than on its function in thinking. One important function of the act of visualizing may be to "reveal" or make available knowledge about properties, such as the shapes of objects or motion constraints, that is otherwise tacit or implicit. This can be an important strategy in reasoning and problem solving. For example, consider the problem in Figure 2 in which people had to judge the angle at which water would pour out of two glasses, a fat glass and a thin glass. People perform poorly on this task when they answer on the basis of explicit knowledge. However, when they close their eyes and rotate an imaginary glass to indicate the answer, they are almost always correctly judge that water will start pouring out of the fat glass first, and there is no systematic relationship between their answers when they are and are not instructed to use imagery. Visualizing the situation seems to reveal knowledge of which the person was otherwise unaware. The same process seems to occur when we take a mental walk to "discover" how many windows are in our house, or when constructing a visual image reveals that grass is lighter in color than pine trees. Clearly the knowledge was there to begin with, but because the knowledge is linked to our perceptual experience, it may take an act of mental imagery to make it available to our thought processes.

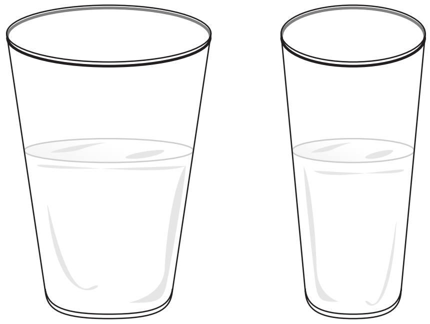


Figure 2 The Water Pouring Problem

Source: Schwartz. D. L. (1999). Physical imagery: Kinematics vs. dynamic models. *Cognitive Psychology, 38*(3), 433–464.

Although there is still much debate about the form and function of mental imagery, there seems to be no doubt that visual imagery and visual perception share mechanisms and brain structures, so that as our understanding of the visual system increases, we will be better able to solve the mystery of visual imagery.

Mary Hegarty

See also Auditory Imagery; Brain Imaging; Consciousness; Neuropsychology of Perception; Olfactory Imagery; Spatial Memory; Transcranial Magnetic Stimulation; Visual System Structure

Further Readings

Finke, R. M. (1989). *Principles of mental imagery.* Cambridge: MIT Press.

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery.* New York: Oxford University Press.

Pani, J. R. (1996). Mental imagery as the adaptionist views it. *Consciousness and Cognition, 5*(3), 288–326.

Pylyshyn, Z. W. (2003). *Seeing and visualizing: It's not what you think*. Cambridge: MIT Press.

Schwartz. D. L. (1999). Physical imagery: Kinematics vs. dynamic models. *Cognitive Psychology, 38*(3), 433–464.

Shepard R. M., & Cooper, L. (1982). *Mental images and their transformations*. Cambridge: MIT Press.

Tye, M. (1991). *The imagery debate*. Cambridge: MIT Press.

VISUAL LIGHT- AND DARK-ADAPTATION

The human visual system works effectively over an enormous range of light intensities from starlight to bright sunlight. This is made possible by adjustments of its operating level to match the ambient illumination: *Light adaptation* is an adjustment to higher levels of illumination, and *dark adaptation* is an adjustment to lower levels. This entry covers the processes that mediate adaptation, adaptation and visual processing, scotopic recovery processes, color, anomalies of dark adaptation, light adaptation, and electrophysiology.

Processes That Mediate Adaptation

The pupil opens wide at night and thereby lets in up to 10 times more light, but this is only a small part of the entire adaptive range. (This effect is greater in species with slit pupils.) Adaptation in humans is instead controlled primarily within the eye, by the retina. The retina is duplex, possessing both a *photopic* system dependent on *cone* receptors that operates in daylight and a *scotopic* system dependent on *rod* receptors that operates in moonlight and starlight, when the photopic system is insensitive. In twilight, the intermediate or *mesopic* range, both receptor classes are active. In daylight, the rods saturate—their responses become constant, independent of the visual stimulus—and vision becomes entirely photopic because the cones are protected from saturation by light adaptation. Light adaptation in bright sunlight is photochemical—many light-sensitive visual pigment molecules in the cones are bleached, becoming transparent, and, with fewer active pigment molecules to pick up light, saturation is avoided. At lower light levels bleaching does not occur, and adaptation in both rods and cones is neural, in part reflecting modifications of synaptic transmission.

Adaptation and Visual Processing

The transition from photopic to scotopic vision is not just a matter of changing receptors, but also of a fundamental change in the nature of visual processing. In sunlight, the photopic system does not attempt to record every photon—there are so many that this would be pointless—but rather it encodes *contrast,* the intensity of light at one location relative to that from an adjacent location. Thus surfaces of objects are encoded as darker (negative contrast) or lighter (positive) than their surrounds. Encoding contrast implies discounting the ambient illumination, so a sunny sky looks only a little darker when the clouds roll in, even though the ambient light level may decrease a thousand times. In starlight, there are few ambient photons, and the ability of the fully dark-adapted scotopic system to detect individual photons becomes critical. Negative contrast no longer exists as the surroundings are dark; processing is devoted instead to small increments in light reflected from dim surfaces. In moonlight, contrasts begin to be encoded as well as increments.

The time course of dark adaptation reveals both the changeover from photopic to scotopic vision and a progressive improvement of sensitivity within each system. The solid symbols in Figure 1 plot psychophysical (behavioral) measurements of visual thresholds obtained after prior exposure of the eye to a bright adaptation light, which was extinguished at time zero. Each threshold indicates the intensity of a briefly flashed test light just needed for an observer to report seeing the test in the dark. Test flashes whose intensities are below the threshold are invisible; those above it are visible. As dark adaptation progresses, the plot reveals

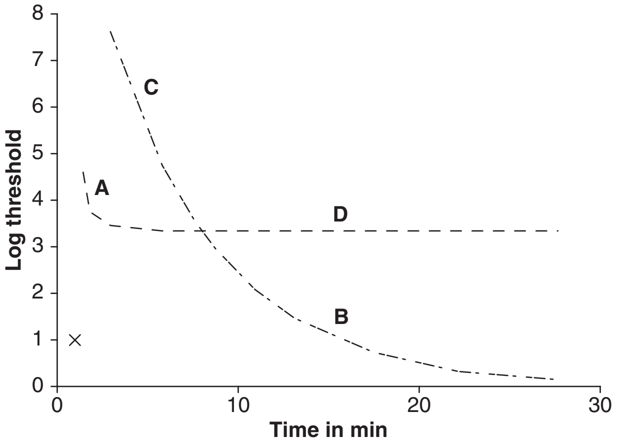


Figure 1 Dark Adaptation

Notes: The solid symbols show thresholds for detection in logarithmic units for observer R. A. Weale, plotted against the length of time in min that his eye was in darkness. Thresholds are photopic (A) or scotopic (B), as determined by the most sensitive system. The dashed line (D) is an extrapolation of the photopic limb. The dot-dash line (C) is an idealized scotopic recovery curve (time constant –0.162).

a descending curve, as decreasing amounts of light are needed to just see the test flash, which corresponds to an increase in visual sensitivity of 4.6 logarithmic units (40,000 times). The curve has two limbs: (A) in the first 2 min, detection thresholds drop rapidly and then plateau; later on, (B) thresholds again drop, reaching a final level after about 30 min. These progressive drops in threshold correspond to increases in visual sensitivity in each system, the changeover from cones to rods or *rod-cone break* occurring at 7 min in Figure 1. Evidence that the break indeed corresponds to a change of system is provided by a shift (the *Purkinje shift*) in the peak spectral sensitivity of the eye from 555 nm (the photopic maximum) before the break to 500 nm, the maximum sensitivity of the rod photo-sensitive pigment, rhodopsin, after the break. Test lights whose intensities are just above the detection threshold appear colored if flashed before the break, indicating cone function, and colorless after it, indicating rod function. The thresholds in Figure 1 were obtained with a middle-wavelength test flash delivered to the *periphery*—a retinal area outside the fovea that contains both rods and cones. When dark adaptation is measured with long-wavelength tests invisible to the rods, or with test flashes of any color presented to the rod-free fovea, there is only a photopic limb, as indicated by the dashed line (D). When dark adaptation is measured in a rod monochromat, a rare individual without cones, the beginning of the scotopic recovery curve (C) is uncovered. The two limbs of the dark adaptation curve shown in Figure 1 can each be shifted vertically by altering the color or brightness of the preadaptation light and by varying the wavelength, size, retinal location, and duration of the test flash, all factors that govern visual sensitivity. If the two limbs do appear, their time courses will be as shown, indicating that the recovery curve for each receptor class is independent of the state of adaptation of the other.

Scotopic Recovery Processes

Independent recovery implies that it is possible to study the biochemical and physiological processes that account for adaptation separately for each type of receptor, rod and cone, without concern for complicating interactions. Recovery processes have been well characterized for the rods, which capture photons via an interaction with the visual pigment molecule, rhodopsin. Capture causes a small component of the rhodopsin molecule, *retinal,* to separate almost instantly from the *opsin;* recovery, which takes time, requires regenerating the rhodopsin. The eye becomes increasingly sensitive during dark adaptation as the metarhodopsins and free opsin, which remain in the receptors when the eye is plunged into darkness, are progressively removed. Removal of these substances, along with regeneration of rhodopsin, help to explain the scotopic limb of the threshold recovery curve (B) seen in Figure 1. Separation of retinal from opsin also triggers a G protein enzyme cascade, which massively amplifies each tiny photon-capture event, permitting signaling molecules to modify the rod cell membrane potential and ultimately communicate the presence of light to the brain.

After extended dark adaptation, the eye becomes so sensitive that individual photons of light can be detected.

Color

The procedures of light and dark adaptation are useful for investigating color vision. If, for example, thresholds are measured in the dark for monochromatic test flashes following light adaptation to monochromatic backgrounds, the photopic curve shown in Figure 1 may exhibit various limbs, depending upon the particular wavelengths and adapting intensities used. For example, following a red light adaptation, detection of a long-wavelength test flash in the dark may be mediated by middle-wavelength sensitive (M) cones at first, but later on by long-wave (L) cones. By measuring thresholds for a wide range of test and adaptation wavelengths, it is possible to determine the spectral sensitivities of these limbs and thus to define the color processes they represent.

Anomalies of Dark Adaptation

The recovery of threshold shown in Figure 1 is not always obtained. Photopic thresholds may fall initially, as shown in Figure 1(a), but after several minutes in the dark, start to rise again; this is the Lie effect, obtained if the observer must detect the color, not just the presence, of the test flash. In some cases, even detection thresholds may rise immediately after the eye is plunged into darkness, remaining elevated for many tens of seconds. A case of especial interest is that of *transient tritanopia,* the temporary blindness to short-wave test flashes that can occur after turning off a yellow adapting field. Yellow fields light adapt the M and L cones, but not the S-cones that mediate detection of the short-wave tests, so the elevation of threshold cannot occur in the cones themselves but must occur at a neural site proximal to the cones. Similar violations of the typical course of dark adaptation can also occur for red and green tests and for fast flicker (*transient lumanopia*). For both flicker and hue, these violations depend strongly on the wavelength composition and luminance of the adaptation light. They have not yet received explanation at the physiological level, although some of the data are compatible with an opponent-process account of color vision. Thus, the recovery of vision shown in Figure 1 is not always representative.

Light Adaptation

One measure of the extent of light adaptation is provided by the increment threshold, the threshold for detecting a test light flashed on a steady field of light to which the observer has adapted for several minutes. Figure 2 plots the increment threshold as it depends on the intensity of the field. At low levels, the field has no effect, but as the field intensity is raised, the threshold rises, at first gradually and then more rapidly. At higher levels there is a plateau followed by a second rise. As in dark adaptation, the two limbs of the plot reflect the operation of the rod and cone systems. When the eye is adapted to low levels, the rod system is the most sensitive, but at higher levels, scotopic sensitivity drops below photopic sensitivity, as indicated by the upper dashed line, and cones take over detection. The positions of the limbs depend on the size, duration, retinal location, and color of the stimuli, just as in dark adaptation. The example just examined is typical of yellow tests presented against a blue-green background in the parafoveal region. A blue test, which favors the scotopic system, presented against a red adaptation, which desensitizes the photopic system, can produce an increment threshold curve in which the scotopic system is completely uncovered, as indicated by a continuation of the upper dashed line in Figure 2.

Theoretical accounts of the incremental threshold curve invoke an intrinsic activity or “dark light” within each receptor system that combines with the excitation produced by the field to raise the threshold. Dark light is responsible for the absolute thresholds of the scotopic (A) and photopic systems (B). Dark adaptation and incremental threshold curves can be related to one another assuming that the dark light persists after the eye is plunged into darkness. The dark

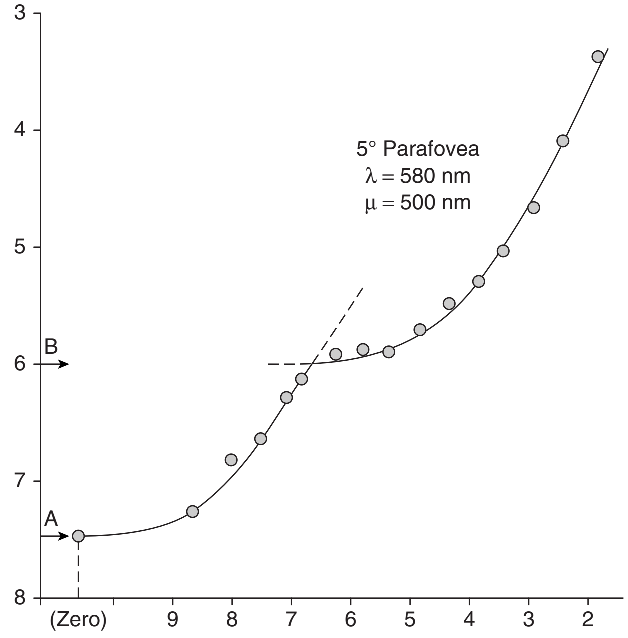


Figure 2 Light Adaptation

Source: Reprinted from Wyszecki and Stiles, 1982.

Notes: Open symbols are incremental thresholds for a brief test flash, plotted against the intensity of the light-adapting field. The test was yellow (580 nm) and the field was blue-green (500 nm) in order to expose both scotopic (lower solid curve) and photopic (upper solid curve) limbs of the threshold versus intensity curves. Axes are in logarithmic units. "Zero" refers to no background light, where the scotopic absolute threshold (A) is obtained. Dotted lines represent continuations of the limbs measured with other wavelength combinations; (B) indicates the photopic absolute threshold. The measured threshold is determined by the most sensitive system.

light acts like a field of light but, being stabilized on the retina, is invisible. Dark light may in part be accounted for by metarhodopsin and free opsin.

Electrophysiology

The phenomena of dark and light adaptation have also been investigated electrophysiologically using single-unit recording, single-cell recording, and electroretinography in various animal species, or just electroretinography in humans, with the goal of explaining the changes that occur with extended exposure to light or darkness. Retinal potentials often show a progressive increase in responsiveness that advances with dark adaptation. Thus, they are useful because they provide objective information regarding the visual process away from the threshold conditions to which psychophysics is limited. The electrophysiological changes seen with light and dark adaptation resemble those seen psychophysically, although exact comparisons are difficult as the stimulus conditions differ. Nevertheless, even the reduction of blue sensitivity found psychophysically at the start of dark adaptation (transient tritanopia) has been seen in the electroretinogram. Thus, although a complete description of light and dark adaptation remains some way off, various aspects of these processes have yielded to scrutiny at the physiological and behavioral levels.

Adam J. Reeves

See also Visual Receptors and Transduction

Further Readings

Arden, G. G., & Weale, R. A. (1954). Nervous mechanisms and dark adaptation. *Journal of Physiology, 125*(3), 417–421.

Hecht, S., Haig, C., & Chase, A. M. (1937). The influence of light adaptation on subsequent dark adaptation of the eye. *Journal of General Physiology, 20*(6), 831–850.

Hess, R. F., Nordby, K., & Sharpe, L. T. (1990). *Night vision.* Cambridge, UK: Cambridge University Press.

Kaiser, P. K., & Boynton, R. M. (1996). *Human color vision* (2nd ed.). Washington, DC: Optical Society of America.

Lamb, T. D., & Pugh, E. N., Jr. (2004). Dark adaptation and the retinoid cycle of vision. *Progress in Retinal and Eye Research, 23*(3), 307–380.

Lie, I. (1963). Dark adaptation and the photochromatic interval. *Documenta Ophthalmologica, 17*(1), 411–510.

Mollon, J. (1982). Color vision. *Annual Review of Psychology, 33,* 41–85.

Naarendorp, F., Sato, Y., Cajdric, A., & Hubbard, N. (2001). Absolute and relative sensitivity of the scotopic system of the rat: Electroretinography and behavior. *Visual Neuroscience, 18*(4), 641–656.

Reeves, A., & Wu, S. (1997). Transient lumanopia: The invisibility of flicker in early dark adaptation. *Journal* *of the Optical Society of America, A., 14*(9), 2509–2516.

Rodieck, R. W. (1998). *The first steps in seeing.* Sunderland, MA: Sinauer Associates.

Schnapf, J. L., & Schneewies, D. M. (2001). Electrophysiology of cone photoreceptors in the primate retina. In K. R. Gegenfurtner & L. T. Scharpe (Eds.), *Color vision: From genes to perception* (pp. 103–112). Cambridge, UK: Cambridge University Press.

Werner, A., Smith, V., Pokorny, J., Kremers, J., & Greenlee, M. (2005). Psychophysical correlates of identified physiological processes in the human visual system. In J. Kremers (Ed.), *The primate visual system* (pp. 311–358). New York: Wiley & Sons.

Wyszecki, G., & Stiles, W. S. (1982). *Color science, concepts and methods, quantitative data and formulae* (2nd ed.). New York: Wiley & Sons.

VISUALLY GUIDED ACTIONS

Humans interact with objects that have diverse sizes, shapes, material properties, and locations, so the control of action must be intricately intertwined with perceptual systems to ensure movement success. Vision provides precise spatial and temporal information about the objects that are the targets of our actions. This entry discusses the role of perception in visually guided action, and how it differs from other forms of perception.

Perception and Action

From an evolutionary standpoint, the organization of perceptual systems must be considered within the context of movement control because natural selection operates at the level of overt behavior: A perceptual system that merely leads to an internal representation or awareness of the external environment without having any ability to influence body movement would not have evolved. Whereas all perceptual functions need not lead to movement, some of them certainly must. The present entry focuses on voluntary movements, rather than reflexive movements.

Action: Voluntary Movement

Whereas reflexive movements are triggered by specific patterns of sensory input, voluntary movements can be elicited whether or not there is an obvious sensory trigger. Voluntary movements have a purpose or a goal, and for this reason they are often referred to as goal-directed movements or actions. The goal of many voluntary movements is to interact with an object in the world, such as a morsel of food or a utensil, to achieve a meaningful outcome, such as satisfying one's hunger, or completing a required task. In object-directed actions, sophisticated perceptual functions are necessary to guide the movement. After all, different actions are required for forks, knives, and spoons, and different grasping actions are required for a dining fork compared with a pitchfork.

Selecting, Planning, and Controlling Action

Every voluntary movement begins with a decision to engage in a particular course of action. For example, when walking down a city street, one may fortuitously encounter a sidewalk vendor selling roasted pine nuts. If one is hungry, a decision may be made to follow a course of action leading to consumption of the delectable product, but if one is not hungry, or perhaps late for an important meeting, one may wisely decide to proceed according to the original walking route. Sophisticated perceptual and cognitive processes are required to handle response selection decisions, in the sense that high-level goals and motivation must be integrated with recognition of the objects in, and the spatial layout of, the current environment.

Once a course of action has been chosen, it must be enacted. Drawing on the foundational work of Robert S. Woodworth, a distinction can be made between the planning of a movement, reflecting those processes leading up to the initiation of the response, and the control of that movement, reflecting those processes responsible for monitoring and adjusting the response while it is being executed. Motor control scientists have long been interested in the different perceptual and motor processes that take place during the planning and control phases of voluntary movement because the two phases are concerned with two rather different problems.

The movement planning phase is concerned with generating the initial motor commands that will bring the actor close to the intended outcome. The role of perception in movement planning is typically conceptualized as a series of sensorimotor transformations in which spatial information about the target object is transformed from a retinal frame of reference to a frame of reference that is centered on the appropriate effector. Movement planning is a feedforward control process, whereby motor commands are sent to the muscles to be executed without modification in response to sensory feedback.

The movement control phase is concerned with adjusting the ongoing action in response to errors or other unexpected events that are detected via sensory feedback during the movement. This phase is typically conceptualized as a sensorimotor feedback loop in which the current state of the body is continuously compared to an ideal or reference state, with adjustments made as necessary.

Prehension

Manual prehension refers to the compound action of reaching toward a target object with the arm while adjusting the digits of the hand to enclose the object, culminating in the manipulation of the object in a functional manner (e.g., reaching to grasp a fork and then using it to stab a particularly juicy spear of broccoli). Following the foundational work of Marc Jeannerod, manual prehension is conceptually divided into an *approach phase,* reflecting the movements occurring prior to physical contact with the intended target object, and a *manipulation phase,* occurring after the object has been contacted. The approach phase is guided by visual information from the target object, the environment, and the grasping limb, and by proprioceptive information arising from the muscles, tendons, and joints engaged in the task. Somatosensory information is additionally incorporated into the manipulation phase as the fingers make contact with the target object.

Jeannerod's psychophysical and neuropsychological studies led to the conclusion that the approach phase consists of two parallel, but coordinated, control processes. The movement of the hand and arm toward the target is referred to as the *reaching component,* whereas the opening and closing of the grasping digits is referred to as the *grasping component.* Jeannerod proposed that the reach and grasp were controlled by different processes because arm and finger kinematics were sensitive to different object features. However, because the timing of the reach and the grasp components were tightly coupled, it was proposed that the two control processes are coordinated by a central controller. Much research in the area of voluntary sensorimotor control has used prehension as a model task, even though it is only one of many common visually guided actions.

Distinct Operations for Perception and Action

It is often assumed that the perceptual mechanisms that lie behind the control of voluntary action are the same as those that lie behind one's conscious perceptual interpretation of the object's properties. To propose otherwise would seem to fly in the face of conventional wisdom about the voluntary nature of human behavior. However, as David Milner and Melvyn Goodale discuss in detail, there is considerable evidence that the perceptual processes that underlie conscious awareness are quite distinct from those that are used to guide object-directed action.

Human Neuropsychology

If it were the case that object-directed actions were guided by conscious perceptual experience, then one would predict that individuals who had lost the ability to perceive visual objects should be unable to make accurate actions toward those objects. At some level, of course, this is true. For example, if an individual were blinded because his or her retinas were destroyed, there is no way the individual could detect visual information about an object. As a consequence, the individual would be unable to visually perceive any object and would also be unable to use vision to guide an action toward that object. However, visual perception can be disrupted because of damage anywhere along the pathway connecting the retinas to higher cortical visual centers. For example, damage to the primary visual cortex in the occipital lobe can lead to a phenomenon known as cortical blindness, or blindsight, in which the individual experiences the sensation of being unable to see objects that are located within the receptive field of the damaged tissue. As far as the individual is aware, he or she is completely blind to the object. However, if asked to point toward the object, the accuracy of the pointing movement is higher than chance would predict. That is, even though the individual reports that no object is present in the surrounding world, and that it is preposterous to even attempt to point to an object that is not present, the direction of the pointing movement actually reflects some knowledge of the object's location. Clearly, some visual information about the object is able to reach the systems in charge of controlling pointing movements, even though that same information is unable to engage the systems responsible for conscious perception.

Whereas blindsight affects all aspects of visual perception in a particular region of space, other forms of brain damage can affect more specific perceptual abilities, such as shape perception. In a particularly poignant example of such a disorder, a young woman known by the initials D. F. suffered an episode of carbon monoxide poisoning that left her with damage in a region of the visual cortex known as the lateral occipital complex. This region of tissue lies in the ventral visual stream, which plays a key role in many perceptual aspects of vision. D. F.'s injury left her with the inability to recognize objects from their visual appearance. Because her deficit is most pronounced for simple black-and-white line drawings, it appears quite selective for the perception of shape or form and is referred to as visual form agnosia. Despite her inability to perceive the shapes of objects, D. F. is nevertheless quite adept at picking them up with her hands by shaping her hand posture quite nicely to fit the shape of the object. In other words, D. F.'s abnormal shape perception seems to have no adverse effect on her ability to use shape information to control her actions. A popular account of this interesting case holds that visual perception and visual control of action are mediated by different parts of the visual brain; visual perception is mediated by the ventral stream leading from the occipital lobe to the temporal lobe, whereas the visual control of action is mediated by the dorsal stream leading from the occipital lobe to the parietal lobe.

Visual Illusions

The evidence previously presented strongly implicates separate mechanisms in the brain for visual perception and for visual control of action. Because this evidence comes from cases in which the brain has been damaged, it is not clear if the distinction between perception and action has any implication for function in the intact brain. A study by Salvatore Aglioti and colleagues provided some tantalizing support for a functional dissociation between visual perception and visually guided action in the intact brain. Using a size-contrast visual illusion, it was shown that participants misperceived the sizes of circular disks but nevertheless scaled the opening of their grasping hand to the true sizes of those very same disks. Several subsequent studies have produced similar results with different types of illusions, different types of perceptual judgments, and a variety of motor actions, but a number of other studies have disputed these findings. Indeed, the question of whether or not actions are resistant to perceptual illusions has generated an extensive body of literature that suggests a single unifying explanation is unlikely. Thus, although the visual illusion paradigm provided early support for the existence of separate mechanisms for perception and action, this simple interpretation cannot explain the diversity of results found in the existing studies.

David A. Westwood

See also Action and Vision; Agnosia: Visual; Reaching and Grasping

Further Readings

Aglioti, S., DeSouza, J. F. X., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology, 5,* 679–685.

Buneo, C. A., & Andersen, R. A. (2006). The posterior parietal cortex: Sensorimotor interface for the planning and online control of visually guided movements. *Neuropsychologia, 44,* 2594–2606.

Elliott, D., Helsen, W. F., & Chua, R. (2001). A century later: Woodworth's (1899) two-component model of goal-directed aiming. *Psychological Bulletin, 127,* 342–357.

Jeannerod, M. (1986). The formation of finger grip during prehension: A cortically mediated visuomotor pattern. *Behavioural Brain Research, 19,* 99–116.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action.* New York: Oxford University Press.

Weiskrantz, L., Warrington, E. K., Sanders, M. D., & Marshall, J. (1974). Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain, 97,* 709–728.

Woodworth, R. S. (1899). The accuracy of voluntary movement. *Psychological Review, 3,* 1–119.

VISUAL MASKING

In *visual masking,* a visual stimulus called the "target" becomes less visible due to interactions with other stimuli, called "masks." The 19th century scientist S. Exner first discovered masking in the visual system. The birth of visual masking was an artifact in one of Exner's studies of consciousness. He had been trying to determine the shortest flash duration necessary for a bar of light to be visible. As a control condition, he presented two identical bars in different places of the visual field and at different times, expecting that they would be perceived as identical in appearance. Exner was surprised to find that this was not, in fact, so. Under certain specific timing conditions, the first bar was rendered invisible by the presentation of the second bar.

In masking experiments, the target and mask can be presented simultaneously, the mask can be presented before the target ("forward masking"), or the mask can be presented after the target ("backward masking"). Figure 1 illustrates backward masking. In this example, the target is a single bar that is visible when presented alone (Figure 1a). The masking stimulus is two bars that could potentially flank the single bar (Figure 1b). When the target stimulus is presented first, and then is immediately followed by the masking stimulus, the target stimulus is rendered invisible (Figure 1c). Because the mask is presented after the target, the effect operates as if the brain's response to the mask somehow catches up to the brain's response to the target so as to inhibit the target's response and suppress the target's visibility. Moreover, when the target and mask are presented simultaneously, there is little or no masking, suggesting that the brain mechanisms that cause masking involve some sort of time-delay system. Thus, masking can occur in some conditions when the target and mask overlap each other spatially and/or temporally, or in other conditions when they do not overlap each other either temporally or spatially. The specific spatiotemporal conditions that lead to masking are important to understand because their neural correlates are critical, presumably, to understanding the neural basis of visibility and perception. Masking can therefore be used to examine the brain's response to the same physical target under varying levels of perceptibility.

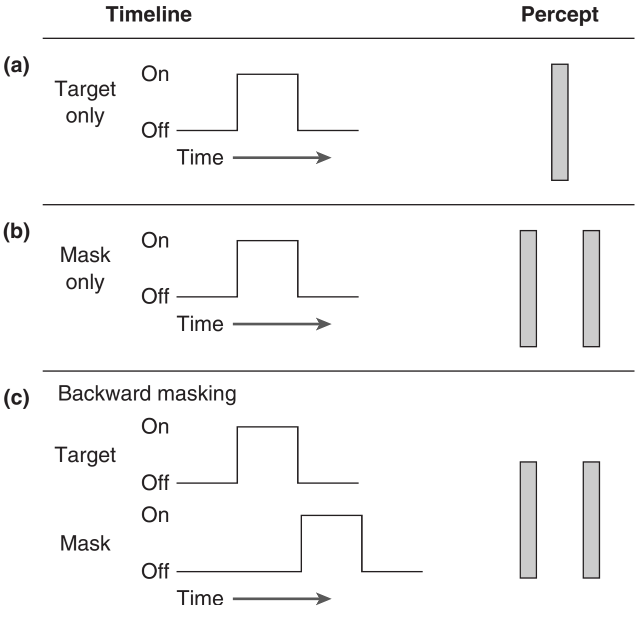


Figure 1 Description of the Perception of a Target and Mask With Respect to Temporal Arrangement

Notes: (a) The target presented alone appears as a single black bar. (b) The mask presented alone. (c) When the target is presented first, followed by presentation of the mask, the target is not visible (it has been masked) so all that is seen is the masking stimulus.

Stephen L. Macknik

See also Vision: Temporal Factors

Further Readings

Breitmeyer, B., & Ogmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision* (2nd ed.). Oxford, UK: Oxford University Press.

Macknik, S. L., & Martinez-Conde, S. (2007). The role of feedback in visual masking and visual processing. *Advances in Cognitive Psychology, 3*(1–2), 125–152.

Macknik, S. L., Martinez-Conde, S., & Haglund, M. M. (2000). The role of spatiotemporal edges in visibility and visual masking. *Proceedings of the National Academy of Sciences USA, 97*(13), 7556–7560.

VISUAL MEMORY

Visual memory can be defined as the maintenance and processing of visual information after the source of that information is no longer available to the visual system. More than a store, visual memory can be considered as an encoding system for storing and retrieving information in a format that is inherently visual (i.e., retaining some of the original properties generated at the moment of sensory or perceptual processing), and therefore different from other formats, such as the verbal one. For example, whereas verbal information is more often encoded with a semantic (expressing the meaning of the verbal information) or phonological code (expressing the sound of the words), visual information is inscribed and retrieved from memory in the form of images and/or visual features. Evidence for the existence of a visual code in memory comes from various sources: the existence of detailed visual memories in nonhuman primates, the ability to discriminate between large numbers of complex abstract fractals that are semantically meaningless; the high precision of stored sensory information, necessarily involving the representation of visual features; the findings that patients with occipital brain lesions (affecting visual cortices) might have a specific visual memory deficit, whereas patients with mediotemporal lesions (affecting "memory" areas) might have spared visual memory abilities; and finally, brain imaging studies showing the involvement of visual areas in visual memory tasks.

However, visual memories do not always have the form of mental images (mental representations of a visual object or a visual scene, accompanied by an experience that significantly resembles the experience of perceiving an object, event, or scene, but that occurs when the relevant object, event, or scene is not actually present to the senses). Their format can range from visual images to more abstract, nonimagistic kinds of representations (although still "visual," that is, coding visual characteristics), depending on the type of visual memory and on the level of processing. As with other memory systems, a taxonomy of visual memory can encompass three different areas: visual sensory memory (VSM), visual short-term memory (VSTM), and visual long-term memory (VLTM). This entry discusses each of these in turn.

Visual Sensory Memory

VSM (also called iconic memory) is an extremely brief type of visual memory (during less than half a second) able to store accurate information about the perceived image. It reflects visual processing at early levels of the visual pathways (between the retina and the visual cortices) and is not influenced by attention of will. It decays gradually and autonomously, although it is subject to interference (masking) due to visual items presented in close temporal proximity (e.g., 100 milliseconds delay) at the same or close locations.

VSM can be divided in two categories: visual persistence and informational persistence. Visual persistence is a picturelike representation that reflects the phenomenonological experience of fading of a visual image that lasts longer than the actual stimulus duration (such as the persistence of a flash of a camera in the dark). Visual persistence is involved in visual integration of successive images, optimal for intervals shorter than 150 milliseconds, and is due to the residual activity of neurons at early stages of visual processing, from photoreceptors in the retina to neurons in the primary visual cortex (V1).

Informational persistence instead pertains to the possibility of extracting visual information after the visual stimulus has been removed, lasts longer than visual persistence, and reflects the decaying of activity of neurons at the level of striate and extrastriate visual cortices.

How much information can be stored in VSM? When observers are asked to report a large amount of information displayed for a brief time (say a

4 × 3 matrix of letters displayed for 100 milliseconds), they are able to report only a small fraction of it (say 4 or 5 letters). George Sperling was able to show that this result is not due to a small capacity of the VSM, but to its brevity: by the time the first four letters are reported, the memory of the other letters has already been lost. In fact, by asking observers to report only the row of letters indicated by a cue appearing just after the letter's presentation (and therefore reducing the limitations due to the time required to report them), observers are able to report nearly all letters of the specified row. Note that all letters have to be stored in VSM in order to perform this task: when the cue indicates the row to be reported, sensory information about the matrix of letters is no longer available. Indeed, it has been estimated that nearly all displayed letters are stored for a short time (few hundred milliseconds) in VSM. Thus, VSM has a large capacity, but it decays rapidly.

What kind of information is stored in iconic memory? Researchers have shown that iconic memory can represent visual features such as location, color, size, and shape. Instead, category cues (such as letter versus digit or vowel versus consonant) needing a higher level of representation cannot be stored (or accessible) in iconic memory.

Visual Short-Term Memory

VSTM is slightly longer lasting than VSM: Visual information can be stored for various seconds in VSTM. When the information stored in VSTM is used for other ongoing cognitive tasks (such as trying to compare two figures representing the same object from different points of view), it is also called visuospatial working memory.

VSTM is intrinsically different from other kinds of short-term memory modalities, such as verbal short-term memory. In fact, whereas strong interference is found in dual tasks involving the same memory modality (verbal or visual short-term memory), there is little or no interference in dual tasks each involving a different type of memory. Moreover, different brain areas are involved in each of the two short-term memory modalities.

The format of visual representation in VSTM is more abstract than that in VSM (does not necessarily involve high-fidelity, picturelike representations), occurs at higher levels along the visual cortical pathways, and it is robust. In fact, it is much less affected than VSM by aspects that influence low-level memory representations (i.e., the position of the stimuli, masking, eye movements, and eye blinks). For these reasons, VSTM seems to have an important role in bridging separate perceptual episodes, thus maintaining a perceptual continuity in scene representation despite these interruptions.

Two different aspects of VSTM have been studied by different researchers. One aspect concerns the precision by which a stimulus or a feature is retained in memory. Authors often refer to this aspect as visual perceptual memory (VPM), which is an implicit (nonconscious) memory that is involved in an object's recognition.

VPM is usually investigated by using simple stimuli that allow direct measurement of simple stimulus features, such as contrast, spatial frequency, orientation, color, and motion. VPM may last for longer than 10 seconds, with an astonishing fidelity for attributes such as spatial frequency or motion, or with a little or pronounced decay respectively for orientation and contrast. The information about simple stimulus attributes seems be maintained in specialized memory stores (one for each attribute) located along the cortical visual pathways.

This kind of memory performance is not related to episodic memory—memory of times, places, associated emotions, and other conception-based knowledge that can be explicitly stated—but it is implicit (can only be found with forced-choice experiments), and its decay is resistant to factors that affect perceptual discrimination, such as an increase in luminance or contrast (i.e., there is no additional impairment due to the memory process).

The second, much-studied aspect of VSTM aims instead to assess the characteristics of memory capacity; that is, how many items or how much information can be retained in VSTM. Two different views dominate the literature on VSTM capacity: According to the "slots" hypothesis, the capacity limits of VSTM are restricted to 4 to 5 representations or slots, independent of their complexity; other authors instead believe that there are no fixed slots and that the capacity of VSTM depends on the amount of information to be stored. According to the latter view, capacity storage would be dependent on the complexity of the stored items more than on the number of items.

The operation that transforms perceptual representations (or VSM) in VSTM is called consolidation; it is fast (20–50 milliseconds) and occurs by means of a limited-capacity central process. Indeed, when two target items are presented at short time intervals among other items, as in a rapid serial visual presentation (RSVP), the lack of consolidation of the second item (due to the capacity limit of the central process, still engaged in consolidating the first target) produces the phenomenon of the attentional blink, that is, a failure in reporting the second target.

Similarly to VSM (and contrary to VLTM), VSTM representations are stored by means of sustained firing of action potentials of neurons in areas of the visual cortices. The posterior parietal cortex has also been found to be involved in VSTM tasks.

Visual Long-Term Memory

VLTM has a remarkably larger storage capacity and a much longer duration than VSTM. Learning of 10,000 pictures can lead to the impressive performance of 84% correct responses in test trials, but this can be attributable to the dual coding (both visual and verbal) of pictures, not just to VLTM. The functions of VLTM can range from object and scene categorization to the use of contextual information about the visual scene in visual search (e.g., presenting background elements having the same features and spatial position can facilitate target detection or discrimination), or the formation of episodic memory representations. It is disputed whether VLTM is formed through a consolidation of VSTM, or whether VLTM is formed via an independent process, directly from perceptual or VSM representations. It is clear, however, that VLTM representations have a similar format to those maintained in VSTM, and its capacity is not affected by the retention of hundreds of visual objects. VPM also seem to contribute to VLTM: The long-term (one week) memory for large numbers of pictures is sensitive to the presence of minor details (usually stored by VPM, not by VSTM) that were actually not noted during the study phase.

Both VSTM and VLTM are largely involved in the representation of visual scenes and in maintaining a stable representation across shifts of eyes and attention within the scene. In particular, thanks to its exceedingly large capacity and the ability to hold information for a long time, the involvement of VLTM occurs when visual memory representations are retained across multiple eye fixations.

As for VSTM, the representation of VLTM is believed to be stored in the same functionally specialized visual cortical areas responsible for the processing of visual attributes. However, the encoding of VLTM memory is not based on sustained neural firing, but upon longer-lasting structural synaptic changes that modify the pattern and the strength of connections between neurons.

Gianluca Campana

See also Attention and Memory; Change Detection; Object Persistence; Perceptual Learning; Priming; Rapid Serial Visual Presentation; Spatial Memory; Visual Imagery; Visual Masking; Visual Scene Perception

Further Readings

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

Campana, G., Cowey, A., & Walsh, V. (2002). Priming of motion direction and area V5/MT: A test of perceptual memory. *Cerebral Cortex, 1,* 663–669.

Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 845–866.

Hollingworth, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 519–537.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory & Cognition, 2,* 683–702.

Luck, S. J., & Hollingworth, A. (2008). *Visual memory.* New York: Oxford University Press.

Magnussen, S. (2000). Low level memory processes in vision. *Trends in Neuroscience, 23,* 247–251.

Nikolic, D., & Singer, W. (2007). Creation of visual long-term memory. *Perception & Psychophysics, 69,* 904–912.

VISUAL PERCEPTION

See Vision

VISUAL PROCESSING: EXTRASTRIATE CORTEX

Despite the ease with which we identify objects in complex environments, the computation of meaningful global forms from local image features on the retina is a challenging task for the visual system. The *extrastriate cortex,* all of the cortex anterior to the primary visual cortex, is implicated in supporting visual mechanisms that convert signals about basic visual features (e.g., position and orientation) to global form percepts in the human brain. In particular, a network of visual areas with selectivity for features of increasing complexity has been implicated in this task: Local image features (e.g., position and orientation) are processed in the primary visual cortex, whereas complex shapes and object categories (faces, bodies, and places) are represented toward the end of the visual pathway in the temporal cortex. This entry reviews evidence from physiology and neuroimaging about the functional role of extrastriate visual areas in the processing of properties critical for visual recognition: object form, depth structure, and motion. This current evidence suggests that the visual system employs a code of increasing efficiency across stages of processing that is advantageous in many respects: It reduces redundancy in the sensory input by integrating information about basic visual features into coherent percepts, provides the building blocks for the representation of complex shapes and biologically important object categories (e.g., faces), and supports fast read out of task-relevant information at different levels of shape description. As such, this code in the extrastriate visual cortex comprises a fundamental computational principle for the analysis of sensory input and is critical for successful detection and recognition of targets in cluttered environments.

Form Processing

How are these neuronal representations that support object recognition constructed in the brain? A network of occipital and temporal areas with selectivity for features of increasing complexity has been implicated in this task. (See color insert, Figure 30.) In particular, neurons in V1 and V2 have been suggested to compute local orientation signals and integrate edges into contours. However, global integration of multiple orientation signals has been attributed to higher occipital regions (e.g., V4) where neurons show selectivity for higher-order features of moderate complexity (e.g., curvature, angles) that define shape parts. Finally, information about object parts is converted into representations of complex shapes (multipart configurations) at the posterior inferior temporal cortex. These shape configurations provide the basis for object recognition at the anterior inferotemporal cortex where neurons selective for entire objects have been identified.

Recently, combined monkey and human functional magnetic resonance imaging (fMRI) studies showed that the perception of global shapes involves both early (retinotopic) and higher (occipitotemporal) visual areas that may integrate local elements to global shapes at different spatial scales. However, unlike neurons in early visual areas that integrate information about local edges and configurations, neural populations in the inferotemporal cortex represent the perceived global form of objects. Importantly, neural representations for global shapes and object categories are characterized by three main principles. First, representations for a large range of object categories (animate and inanimate) have been shown to be distributed across neural populations in the inferotemporal cortex, but specialized modules for the processing of faces, bodies, and places have also been identified. Second, object representations show a retinotopic organization and a center-periphery organization; that is, objects that entail central-vision processing (e.g., faces and words) show center-biased representations (i.e., stronger activation in foveal regions), while objects that entail integration across the visual field (e.g., places) show periphery-biased representations (i.e., stronger activation in peripheral regions). Third, unlike earlier visual areas, patterns of neuronal activity in the inferotemporal cortex contain information about object identity independent of image changes (e.g., object position and size) or cues that define the shape (luminance, texture, and motion). Recent studies suggest a progression of visual object representations from posterior temporal regions that contain information about local image properties (e.g., position and size) to anterior regions that show a higher degree of tolerance to changes to these image properties. Computational models propose that the brain builds robust object representations using neuronal connections that group together neurons tuned to similar image features across image transformations. Recent neurophysiological studies showing that temporal cortex neurons with high object selectivity have low invariance (i.e., tolerance to changes in image properties) suggest that these connections between neurons selective for similar features are critical for the binding of feature configurations and the robust representation of object identity independent of image changes. These findings suggest that although individual neurons contain highly selective information for image features that is critical for interactions with objects (e.g., precise information about position, size, and orientation), connections across neural populations may support object recognition across image changes.

Depth Processing

Our perception of the world's three-dimensional structure is critical for object recognition, navigation, and planning actions. To achieve it, the brain combines different types of visual information (binocular and monocular) about depth structure. The horizontal separation of the eyes provides a powerful cue to depth in the form of binocular disparity. Humans are exquisitely sensitive to disparity and exploit it for multiple purposes (e.g., breaking camouflage, perceiving albedo, and grasping objects). Computationally, different processing stages are required to extract useful information from the images registered by the two eyes. Although neurons sensitive to disparity are found throughout the visual cortex, understanding the circuits supporting different stages of disparity computation and the functional utility of these computations remains an open challenge.

Neurophysiological studies have revealed selectivity for binocular disparity at multiple levels of the visual hierarchy in the monkey brain from early visual areas to object and motion selective areas and the parietal cortex. Imaging studies have identified multiple human brain areas in the visual, temporal, and parietal cortex that show stronger activations for stimuli defined by binocular or monocular depth cues than for 2-D versions of these stimuli. In particular, areas V3A and V3B/KO have been implicated in the analysis of disparity-defined surfaces and boundaries. Further, several recent studies suggest that areas involved in disparity processing, primarily in the temporal and parietal cortex, are also engaged in the processing of monocular cues to depth (e.g., texture, motion, and shading) and the combination of binocular and monocular cues for depth perception.

Motion Processing

Processing of motion information has been primarily attributed to MT/V5, an area known to have neurons selective for the direction and speed of motion. Human MT/V5 is localized in the ascending branch of the inferior temporal sulcus (ITS) and corresponds to the macaque MT/V5 but also to several of its satellites (e.g., MST and FST). Beyond MT/V5, other extrastriate areas have been proposed to be involved in the processing of different types of motion. For example, the processing of relative motion in kinetic boundaries (i.e., displays in which random dots move in opposite directions in alternate stripes) is attributed toV3B/KO (kinetic occipital area) located between V3/V3A and MT/V5. Further, processing of more complex types of motion (i.e., coherent motion and structure from motion) is thought to engage a network of cortical regions in both the ventral and the dorsal stream. Finally, recent imaging studies in both humans and monkeys show responses to 3-D shape from motion in ventral and dorsal parietal regions along the IPS that may relate to the visuomotor control of actions (e.g., human-object/tool interactions, direction of heading). Interestingly, these representations of 3-D shape from motion are more extensive in human than in monkey IPS, potentially due to the more extensive tool use in humans compared to monkeys.

Integrating Form, Depth, and Motion

Successful interaction with the complex and dynamic environments we inhabit requires that the visual system integrates information about the form (shape features and 3-D structure) and motion of objects and actors into dynamic perceptual events. For example, identifying a friend in the crowd entails integrating facial and body features with personal movement styles. Clear evidence that form and motion processing interact comes from psychophysical studies. For example, motion facilitates the detection of shapes embedded in cluttered backgrounds, the invariant representation of objects (i.e., independent of changes in the object position or orientation), and the discrimination and identification of faces. On the other hand, form influences the perception of motion paths. For example, the biomechanical structure of the human body constrains the perception of motion paths (e.g., joints are perceived to move around but not through the human body).

A striking example of form and motion interaction is our ability to infer motion from still images depicted in paintings, photographs, or cartoons. We readily recognize whether an animal, person, or object within a photograph was moving or standing still at the moment a photograph was taken. Photographers, painters, sculptures, and cartoonists can successfully convey motion information, even though no real motion is present in their work. Cartoonists, for instance, use the inanimate form cue of speed lines to suggest motion in an effective manner. Motion in art is often implied by animate form cues. In the case of human figures, these cues include body posture, articulation of arms and legs, and the overall imbalance of the body.

This type of animate implied motion shows how form can lead to a "sense" of motion, and the study of implied motion aims to uncover how the brain uses form cues to generate the percept of motion. In particular, human fMRI studies have shown that pictures with implied motion (e.g., a snapshot of an athlete running or a cup falling off a shelf) evoke stronger responses in MT/V5 than pictures without implied motion. Recent studies using magnetoencephalography (MEG) and visual evoked potentials (VEP) in humans demonstrate a clear interaction between motion and form signals in the visual cortex. In particular, it was shown that the amplitude of an implied motion response was significantly reduced after adaptation to real motion in the same direction as the implied motion direction. Consistent with physiology findings, adaptation indicates that form and motion are not only processed in the same cortical areas, but that responses to implied motion arise from direction-selective neurons that are similarly tuned for real and animate implied motion directions.

Finally, it is of interest to note that experience and learning have been suggested to play an important role in the optimization of visual functions that mediate form and motion perception. Encoding of primitive features may relate to evolutionary and developmental plasticity mechanisms for shaping object selectivity. At a shorter time scale, experience through everyday interactions with objects and exposure to dynamic object transformations may shape the tuning of neural populations. It is possible that similar mechanisms may mediate this long- and short-term optimization of the neural code in the visual cortex through evolution, development, and experience and tune the visual system to the fundamental principles of form and motion binding for object recognition and actions.

Zoe Kourtzi

See also Face Perception: Physiological; Motion Perception: Physiological; Neural Representation/Coding; Object Perception: Physiology; Spatial Layout Perception, Neural; Vision; Visual Processing: Primary Visual Cortex

Further Readings

Brincat, S. L., & Connor, C. E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature Neuroscience, 7,* 880–886.

Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT. *Annual Review of Neuroscience, 28,* 157–189.

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience, 27,* 649–677.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293,* 2425–2430.

Kourtzi, Z., & DiCarlo, J. J. (2006). Learning and neural plasticity in visual object recognition. *Current Opinion in Neurobiology, 16,* 152–158.

Kourtzi, Z., Krekelberg, B., & van Wezel, R. J. (2008). Linking form and motion in the primate brain. *Trends in Cognitive Sciences, 12,* 230–236.

Orban, G. A., Janssen, P., & Vogels, R. (2006). Extracting 3D structure from disparity. *Trends in Neurosciences, 29,* 466–473.

Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience, 8,* 379–391.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience, 19,* 109–139.

Welchman, A. E., Deubelius, A., Conrad, V., Bulthoff, H. H., & Kourtzi, Z. (2005). 3D shape perception from combined depth cues in human visual cortex. *Nature Neuroscience, 8,* 820–827.

VISUAL PROCESSING: PRIMARY VISUAL CORTEX

The term *cultural melting pot* describes the phenomenon where people from different cultures around the world come together to live in the same community. While these people strive to adapt to the language and lifestyle of their new residence, they also often preserve much of their native culture allowing for others in their community to learn about their unique background and experiences. By analogy, the first visual area of the cerebral cortex is also a neuronal "melting pot" for visual information arriving through a variety of channels. Within this first visual cortical area, information from these various channels is mixed to some degree to begin establishing coherent representations of the visual environment. But to another degree, information also remains partially segregated in order to retain some of the original characteristics of the input channel. This entry outlines some of the details of information processing in the first visual cortical area, which is called the primary visual cortex, V1, striate cortex, or area 17. V1 contains a myriad of anatomically and functionally distinct neurons participating in specific cortical circuits. These components are exquisitely organized into a functional architecture that is based on the laminar and columnar structure of the cortex.

Cortical Layers, Circuits, and Columns

Like all of the cerebral cortex, V1 is a laminar structure. It is approximately two millimeters thick and consists of six main layers that differ in terms of the organization of their intrinsic and extrinsic connections (color insert, Figure 31). Visual signals from the retina are relayed to V1 via the lateral geniculate nucleus (LGN) of the thalamus. From the retina to the cortex, visual signals are separated into three different channels containing anatomically and physiologically distinct neuronal types, which synapse onto segregated zones within V1. Magnocellular and parvocellular channel inputs target layers 4Cα and 4Cβ, respectively, while koniocellular channel inputs bypass layer 4C and terminate in the cytochrome oxidase rich blobs of layers 2/3. These channels are significant because they have different physiological properties, which are described later. Once visual signals arrive in V1, a rich network of intrinsic connections provides further processing. V1 supplies visual information to other brain regions via a specific organization of extrinsic projections. Targets of V1 outputs include extrastriate cortical areas of the dorsal and ventral streams and subcortical visual structures.

V1 contains two major classes of neurons—excitatory neurons and inhibitory neurons. Excitatory neurons use glutamate as their neurotransmitter and have dendrites studded with tiny protrusions called spines. Excitatory neurons are further classified as pyramidal neurons and spiny stellate neurons. Pyramidal neurons have a pyramid-shaped cell body, an apical dendrite that spans many cortical layers, and basil dendrites that typically remain restricted to the same layer as the cell body. In contrast, spiny stellate neurons have circular cell bodies and lack an apical dendrite. While both types of excitatory neurons have intrinsic axons that target other neurons within V1, pyramidal neurons also often have an extrinsic axon that targets structures outside of V1, including subcortical areas (e.g., the thalamus, superior colliculus, and/or basal ganglia) and extrastriate cortical areas. Inhibitory neurons, the other main class of V1 neurons, use GABA as their neurotransmitter and have smooth dendrites lacking spines. Inhibitory neurons have intrinsic axons that exclusively target neurons within V1.

The majority of intrinsic connections within V1 are made between neurons with cell bodies aligned vertically along a line perpendicular to the cortical surface. This arrangement of connections underlies a fundamental concept of cortical organization, namely the cortical column. Neurons within a cortical column typically share many characteristics in common with each other, such as receptive field location, ocular dominance, and orientation tuning (described later). Cortical columns do not operate in isolation, but as a network of horizontal connections within V1 interconnect columns over distances of up to several millimeters. These horizontal connections likely provide a substrate for a variety of response properties that include effects from stimuli located beyond the classical receptive field of individual neurons.

Emergent Response Properties

Neurons in the retina and LGN typically display concentric and antagonistic center/surround receptive fields (where receptive fields are the areas on the retina that, when stimulated, influence firing of the cell). This type of receptive field is also found among a subset of neurons in the input layers of V1. Importantly, V1 also contains a variety of emergent receptive fields not found in the LGN. Two major types of receptive fields first encountered in V1 are simple cell receptive fields and complex cell receptive fields. Although both types of receptive fields are selective for oriented visual stimuli, the receptive fields of simple cells have alternating on and off subfields, whereas the receptive fields of complex cells have overlapping on and off fields. Based on the progressive complexity of these receptive fields, David Hubel and Torsten Wiesel proposed a straightforward model for their construction. In this model, inputs from several LGN neurons with receptive fields located along a line of visual space (and matching sign, e.g., on or off, center) converge to establish a simple cell. Similarly, inputs from several simple cells with a variety of phases (i.e., on and off subfields are overlapping) converge to establish a complex cell (color insert, Figure 31). Although results from numerous studies provide strong support for this model, results also demonstrate that other intrinsic inputs play key roles in shaping V1 receptive fields.

In addition to the center/surround, simple, and complex receptive fields of V1 neurons, a variety of other receptive field specializations are found. For instance, some neurons selective to stimulus orientation are also selective to stimulus size and/or the direction of stimulus motion. Neurons may also be selective for the chromatic properties of a stimulus and/or the stereoscopic depth of the stimulus. Given the diversity of neuronal circuits in V1, it seems likely that future work may reveal an even greater diversity of receptive fields among V1 neurons.

Functional Architecture

V1 neurons and their parent columns are organized across the cortical surface according to a number of "functional maps." One prominent map is for retinotopy, whereby regions of visual space that excite the cortex shift progressively across the cortex such that neighboring neurons in the cortex respond to neighboring points in visual space. Like other primary sensory areas in the cortex, the retinotopic map in V1 is distorted such that a greater portion of V1 is dedicated to regions of visual space near the fovea (the central retina that contains the highest density of cones and is therefore optimized for acuity). Another prominent map in V1 is for ocular dominance, whereby alternating columns of cortical neurons respond preferentially to visual stimuli presented to one eye or the other. Finally, V1 contains a map for orientation tuning, whereby a full cycle of preferred stimulus orientations is mapped in a pinwheel-like fashion onto adjacent columns (color insert, Figure 31, orientated black bars illustrated on cortical surface). To ensure that all points in visual space are represented by each eye and all orientations, several pinwheels tile across V1 with their centers frequently aligned over individual ocular dominance columns (color insert, Figure 31, gray shading illustrates ocular dominance stripes across cortical surface).

Processing Streams

As mentioned earlier, neurons in the magnocellular, parvocellular and koniocellular channels carry visual inputs to layers 4Cα, 4Cβ, and the layer 2/3 blobs within V1. These inputs to the cortex differ not only in their anatomical properties but also in their physiological properties. Compared to parvocellular LGN neurons, magnocellular neurons have somewhat larger receptive fields and respond better to low-contrast stimuli and fast-moving stimuli. Magnocellular neurons respond to a broad range of wavelengths of light, whereas parvocellular neurons display chromatic opponency (e.g., red on-center/green off-surround). Although much less is known about koniocellular LGN neurons, a growing body of evidence indicates that these neurons supply V1 with information about shorter wavelength (blue) stimuli.

Beyond the layers of V1 that receive direct LGN input, it is much less clear to what extent V1 neurons mix signals from the magnocellular, parvocellular, and koniocellular streams. Some cortical circuits within V1 keep stream-information largely segregated. One good example of a local circuit dominated by the magnocellular pathway is the circuit connecting neurons in layer 4Cα to neurons in layer 4B, which are strongly direction selective and provide output to the extrastriate area MT, known to be involved in motion perception. Circuits connecting the subdivisions of layer 4C to layer 6 may also retain some stream specificity. However, two other prominent layers with projections to the extrastriate cortex, layers 2/3 and 5, contain many neurons that receive mixed inputs from all of the cortical layers. Given evidence that cortical circuits within V1 are diverse in terms of their degree of mixing versus segregation of the magnocellular, parvocellular, and koniocellular channels, it is likely that V1 plays multiple roles in processing early visual information. Some circuits in V1 serve to maintain stream segregation (e.g., circuits targeting the LGN), whereas others serve to combine stream information (e.g., circuits targeting extrastriate cortical areas), making V1 a "melting pot" for incoming visual information.

Farran Briggs and W. Martin Usrey

See also Cortical Organization; Neural Representation/Coding; Visual Processing: Retinal; Visual Processing: Subcortical Mechanisms for Gaze Control; Visual System Structure

Further Readings

Callaway, E. M. (2005). Structure and function of parallel pathways in the primate early visual system. *Journal of Physiology, 566,* 13–19.

Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A.-S., McNamara, J. O., et al. (2008). Central visual pathways. In D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A.-S. LaMantia, J. O. McNamara, et al. (Eds.), *Neuroscience* (4th ed.). Sunderland, MA: Sinaur.

Reid, R. C., & Usrey, W. M. (2008). Vision. In L. R. Squire, J. L. Roberts, N. C. Spitzer, M. J. Zigmond, S. K. McConnell, & F. E. Bloom (Eds.), *Fundamental neuroscience* (3rd ed.). San Diego, CA: Academic Press.

VISUAL PROCESSING: RETINAL

The retina of the eye provides the essential foundation for visual perception. It contains millions of nerve cells that analyze the visual image and transmit signals to the brain. Much like a camera, the eye forms a physical image of an object in the external world on the surface of the retina at the back of the eye. The next essential step is to respond to the image. This is the job of cells known as photoreceptors. The photoreceptors are extremely small—about 1 to 3 microns in diameter. There are about 130 million of these photoreceptors packed side-by-side, much like pixels in a digital camera. There are two classes of photoreceptors. Rods are the receptors for night vision. Cones are the receptors for daylight and color vision. When light falls on a photoreceptor, the photoreceptor responds with an electrical signal—a change in voltage. The emphasis of this entry is on the events that occur beyond the level of the photoreceptors. The overview is this: The electrical signals of the photoreceptors (labeled 1 and 2 in color insert Figure 32) are transmitted to bipolar cells (4). They, in turn, transmit signals to ganglion cells (6).

Ganglion Cells, Receptive Fields, and Distributed Coding

The ganglion cells send their signals by way of the optic nerve to the brain. These signals are absolutely crucial for visual perception. Blindness results if they are lost due to disease or injury. This entry focuses primarily on the signals of these cells. How do they produce the foundation for visual perception? The signals are nerve impulses. A nerve impulse is a change in voltage that is very brief, about 1/1,000 of second. The impulse does not vary in its size, a principle known as the all or none law. As a result, the ganglion cell does not code visual information by the size of its signal. Rather, it codes by the rate at which it responds. This principle, common to most nerve cells, is known as rate coding. For example, the ganglion cell will respond to a dim light with a low rate of response of 5 impulses per second, whereas it will respond to a bright light with a high rate of, for example, 200 impulses per second.

A ganglion cell will only respond to an object that is located in a specific part of the external world. This means that an individual ganglion cell only "sees" a small part of the total visual scene. This small region is known as the receptive field of the cell. Thus, for a small object, for example, a small bird perching on a tree limb, the ganglion cell will respond only if the image of the bird falls within the receptive field. However, other ganglion cells have receptive fields that correspond to other parts of the visual scene. So if the small bird flies across the sky, its image will occupy a succession of positions and a succession of ganglion cells will respond as the image falls in their receptive fields. In essence, this is the way ganglion cells code for the position of objects in the world. Now suppose you look at a large object, for example, an elephant. The image of the elephant will be very large, so an individual ganglion cell with a small receptive field will only respond to a small part of the image. However, other ganglion cells will respond to other small parts of the image. In this way, the object can be coded or represented from the overall activity in many different ganglion cells. This is an example of what is known as distributed coding: The code for the object is based on the particular distribution of activity found across a population of nerve cells. This principle is also known as parallel processing, in analogy with computer design, and is used widely throughout the retina.

Different Types of Ganglion Cells

A fundamental case of distributed coding is the existence of ON and OFF ganglion cells. ON cells respond while a light is turned on. OFF cells respond while a light is turned off and inhibit responding when a light is turned on. In functional terms, this means that the ON cells are excited by light objects, whereas OFF cells are excited by dark objects. This observation lies at the heart of how the retina codes visual information—by the existence and interaction of excitatory and inhibitory signals. For example, when the receptive field of a ganglion cell is studied in detail, it is often found that there are two parts to the field: a central region and a surrounding region. Light in the center excites the cell, whereas light in the surround inhibits the response. The influence of the surround is an example of a process, found in the eye and brain, known as lateral inhibition. Here, signals travel laterally across the retina so that light in the surround inhibits (i.e., opposes) the response to light in the center. The purpose of lateral inhibition, or center-surround antagonism as it is also called, is to enhance the response to edges and local contrast in the visual scene. For example, a small spot of light falling in the center of the receptive field will produce a strong response, whereas a large uniform patch of light will only produce a weak response.

Differences in the size of the receptive field of ganglion cells provide another basic foundation for visual coding. One class of cells, called the M cells, have large receptive fields, whereas others, called P cells, have small fields. The small field P cells provide the basis for our ability to see fine spatial detail, that is, "to read the fine print." The M and P cells also differ importantly in their response over time. When a light is flashed, the P cells respond with impulses as long as the light remains on. The M cells, by contrast, respond strongly when the light first comes on, but thereafter respond weakly or not at all. Hence, the M cells are called "transient cells." In general, the M cells respond strongly only when there is a change in the intensity or movement of an object. This is of great importance for the survival of most animals and humans because usually things that change in the environment are more important than those that are static. Thus, M cells provide one of many examples of the "wisdom of the nervous system." Certain ganglion cells show even more remarkable selectivity: They respond when an object moves in one direction, for example from right to left, but not at all when the same object moves over the same path but in the opposite direction (left to right).

What about color? For at least a century, it has been recognized that there are two primary psychological attributes of color perception: hue and brightness. Therefore, it is fascinating that recent research in neurophysiology has discovered that some ganglion cells code for brightness, whereas others show striking responses to the wavelength of the light. The latter cells are called color-opponent cells because they give opposite responses to specific color pairs. One group of cells, called R/G opponent cells, is excited by red and inhibited by green light. A second group of cells, Y/B cells, is excited by yellow and inhibited by blue light. Nearly 100 years ago, color theorists speculated that there might be such cells somewhere in the retina or brain. The modern discovery of color-opponent cells in the retina thus stands as a landmark in understanding the basis of visual perception.

Bipolar, Horizontal, and Amacrine Cells

How do the responses of ganglion cells come about? The answer, in a nutshell, is that neural circuits and connections of other nerve cells provide input to and thereby shape the ganglion cell response. These other cells, as shown in color insert Figure 32, are the bipolar (blue), horizontal (yellow), and amacrine cells (orange).

The bipolar cells receive their input from rods and cones, then transform the input and send these transformed signals to the ganglion cells. Signal transmission between most retinal cells is accomplished at synapses via neurotransmitter chemicals that pass from one cell to the next. The response of bipolar cells is caused by glutamate, the neurotransmitter that is released by the photoreceptor. One type of bipolar cells, known as ON bipolar cells, is excited by the transmitter, whereas another type, the OFF bipolar cells, is inhibited by the transmitter. Thus, the bipolar cells split the input from the photoreceptors into two parallel channels of opposite sign, ON and OFF, and then by selective connections, provide the input and thus the basis for the existence of ON and OFF ganglion cells discussed earlier.

The bipolar cells also provide the driving force for center-surround antagonism of the ganglion cells because the bipolar cells respond in opposite directions when stimulated by light in the center or surround of their receptive field. The surround inhibits the center due to lateral inhibition that is carried by a network of horizontal cells (3 in color insert Figure 32). In addition, some bipolar cells are extremely sensitive to brightness contrast, whereas others give opponent color responses. For example, one type gives a positive response to red and a negative response to green. While much more is known about bipolar cells, this overview is sufficient to emphasize that their response is shaped by connections with photoreceptors and horizontal cells, and that the bipolars, in turn, shape the response of ganglion cells. Amacrine cells (5, in color insert Figure 32) are another important group of cells that shape the ganglion cell response. They form a network of complex interconnections between themselves, bipolar and ganglion cells. They provide a second major pathway for lateral inhibition and play a major role in shaping the ganglion cell's response to moving objects.

Neural Pathways and Circuits

Many different pathways and neural circuits are used in the retina. They leave indelible marks on visual perception. One important circuit starts with the rod photoreceptors. A large number of rods connect to several bipolar cells that then connect to a ganglion cell. Thus, many rods can influence the response of the ganglion cell. This provides the basis for the principle of spatial summation by which the individual effects of light falling on a large number of rods, and thus over a large area of visual space, are summed at the ganglion cell. Because rods are extremely sensitive to dim light, this circuitry for spatial summation creates a class of ganglion cells that is very sensitive to dim light and thus, is crucial for night vision.

A radically different neural circuit is set up by cones near the center of the retina in a region known as the fovea. The fovea contains very small cone photoreceptors. In a special circuit, just one cone connects to only one bipolar cell and this, in turn, to only one ganglion cell. Because the receptive field of a single cone is extremely small, the receptive field of the ganglion cell is extremely small. This circuit, sometimes called a "private line," is crucial for seeing very small details and thus for our finest pattern vision.

The retina contains millions of nerve cells of many different functional types that are interconnected in a complex of pathways that define a variety of specialized circuits for processing visual information. In this way, the signal that is sent to the brain, an intricate pattern of nerve impulses in space and time, can be varied in an almost infinite number of ways. Neural circuits in the brain then process these signals further to create even more complex and selective coding. This is precisely what is necessary to represent the almost infinite variety of visual objects that we see in the world.

Dwight A. Burkhardt

See also Color Perception: Physiological; Contrast Enhancement at Borders; Eye: Structure and Optics; Lateral Inhibition; Receptive Fields; Retinal Anatomy; Visual Receptors and Transduction

Further Readings

Canadian Institute of Health Research. (2004, January 1). The brain from top to bottom. Retrieved from http://thebrain.mcgill.ca/flash/i/i_02/i_02_cl/i_02_cl_vis/i_02_cl_vis.html

Dowling, J. E. (1987). *The retina: An approachable part of the brain.* Cambridge, MA: Belknap Press.

Goldstein, E. B. (2010). *Sensation and perception* (8th ed.). Belmont, CA: Cengage/Wadsworth.

Kolb, H., Fernandez, E., & Nelson, R. (2001, February). John Moran Eye Center, University of Utah. Web Vision: How the Retina Works. Retrieved from http://webvision.med.utah.edu/

Sterling, P. (2004). How retinal circuits optimize the transfer of visual information. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (Vol. 1, pp. 234–259). Cambridge: MIT Press.

Wassle, H. (2004). Parallel processing in the mammalian retina. *Neuroscience Nature Reviews, 5,* 747–757.

VISUAL PROCESSING: SUBCORTICAL MECHANISMS FOR GAZE CONTROL

For perception to be useful, one needs to actively orient sensory receptors toward the most relevant stimuli in the outside world. In the case of the visual system, the eyes must be oriented so that light reflected from the visual scene falls on photoreceptors within the retina at the back of the eye. In particular, light from the object of greatest interest needs to fall on the fovea. The fovea is the part of the retina with the densest distribution of photoreceptors and thus provides the highest visual acuity. The line between the fovea and the object of regard is known as the *gaze direction* and movements that change gaze direction are called *gaze shifts*. Our brains possess complex motor pathways for shifting and holding gaze, using eye, head, and body movements. Whether these movements are voluntary or involuntary, the final pathways that ultimately control gaze are found in the subcortical nuclei of the brain stem. This entry describes saccadic gaze shifts, eye muscles and motoneurons, primary and additional brain stem premotor circuitry for gaze shifts, other types of eye movement, and higher-level uses of subcortical visual signals.

Saccadic Gaze Shifts

Gaze can be redirected by moving the eyes alone at very high speeds (up to 900° per second [°/s] in humans). These movements are called *saccades,* and they can be made either involuntarily (e.g., to a sudden movement or noise) or voluntarily (e.g., when visually scanning the environment, locating the source of a sound, or reading). Even when we think we are staring directly at an object, the eyes constantly make extremely tiny saccades (as if the eye is jiggling) to prevent the image from fading and to improve visual acuity. Many scientists think that saccades are guided by feedback mechanisms that compare current eye position (i.e., where am I looking now?) and desired eye position (i.e., where do I want to look?). For gaze shifts larger than 150°/s (e.g., while watching a tennis match), gaze is shifted through a coordinated combination of eye and head movements. The eyes, which are lighter and have less inertia, begin to move toward the new visual target first, followed by the heavier head. The head's relative contribution to the overall gaze shift depends on various parameters, including the oculomotor range of the animal (e.g., cats have a small oculomotor range of ±250, and thus move their heads much more than humans, who have an oculomotor range of ±500), and the position of the eyes in the orbits at the start of the gaze shift (i.e., if the eyes are already pointing in the direction of the target, then the head will contribute more). Once the eyes reach the target, the vestibulo-ocular reflex (VOR) is engaged. The VOR causes the eyes to rotate by the same amount as the head, but in the opposite direction, and thus serves to keep the eyes on target while the head catches up. Eye plus head gaze shifts typically end with the eyes nearly centered in the orbits.

Thus, a combined eye plus head gaze shift can be broken down into three components: (1) eye-in-head: the eye movement relative to the head, (2) head-in-space: the head movement relative to space, and (3) eye-in-space: the eye movement relative to space (i.e., gaze, which is the result of 1 + 2). These components are present for all three directions of motion that the eye and head are capable of moving in: (a) horizontal movements cause gaze to be shifted right and left, (b) vertical movements cause gaze to be shifted up and down, and (c) torsional movements cause gaze to be shifted clockwise and counterclockwise around the line of sight (e.g., bringing the right ear down toward the right shoulder involves a clockwise rotation of the head). Horizontal and vertical movements are important for determining gaze direction, whereas torsion is important for eye and head coordination and binocular perception, and is usually reset to zero at the end of each gaze shift (Listing's law). The brain stem mechanisms associated with gaze shifts carefully monitor and control all three directions of motion (a–c) for all three components of gaze (1–3).

Eye Muscles and Motoneurons

Eye position is controlled by three pairs of push-pull muscles. The horizontal recti control horizontal eye position, whereas the vertical recti and oblique muscles each control a combination of vertical and torsional positions. These six muscles are controlled by three motoneuron pools (i.e., the oculomotor, trochlear, and abducens nuclei) in the brain stem. If one records the activity of an ocular motoneuron during a saccade, it shows a temporal pattern of activation made up of two components known as the *pulse* and the *step*. The pulse is characterized by a transient burst of action potentials that act like a velocity command to drive the eyes in a particular direction and amplitude, and thus bring the eyes onto the new target of interest. The step is characterized by a sustained firing rate that is correlated with final eye position. Without the step, elastic forces around the eyes would cause them to roll back toward a central position.

Primary Brain Stem Premotor Circuitry for Gaze Shifts

The pulse for saccades is generated by burst neurons in the paramedian pontine reticular formation (PPRF) for the horizontal component of eye movements and by the rostral interstitial nucleus of the medial longitudinal fasciculus (riMLF) for vertical and torsional components (Figure 1; see also color insert Figure 33). The output of these burst neurons is sent directly to the motoneurons and indirectly to the *neural integrator*. The neural integrator converts the pulse into a step command and then relays it to the motoneurons. The neural integrators for the horizontal component of eye movements are found in the nucleus prepositus hypoglossi (NPH), whereas those for vertical and torsional components are found in the interstitial nucleus of Cajal (INC). The separation of horizontal commands from vertical/torsional commands in brain stem nuclei mimics a similar separation in the eye muscles themselves. Less is known about head control during gaze shifts, but some recent studies suggest that the same brain stem premotor mechanisms produce similar pulse and step signals for the spinal cord motoneurons that innervate the neck muscles.

Burst neurons in the PPRF and riMLF are ultimately under the control of the superior colliculi (SC), which are located at the top of the brain stem, one on each side of the midline. Each SC is divided into a superficial layer that receives visual input directly from the retina, as well as indirectly from the cerebral cortex, and a set of deeper motor layers that produce the saccade command. These layers are arranged such that left visual space is represented in the right SC, and right visual space is represented in the left SC. The motor layers are arranged topographically, so that the location of neural activity on the SC map determines the size and direction of the elicited gaze shift. The front end of the SC is often called the fixation zone because it is active when gaze is fixed on an object and it suppresses other gaze shifts. As one travels toward the back of the SC, activity produces progressively larger gaze shifts with larger and larger head movement

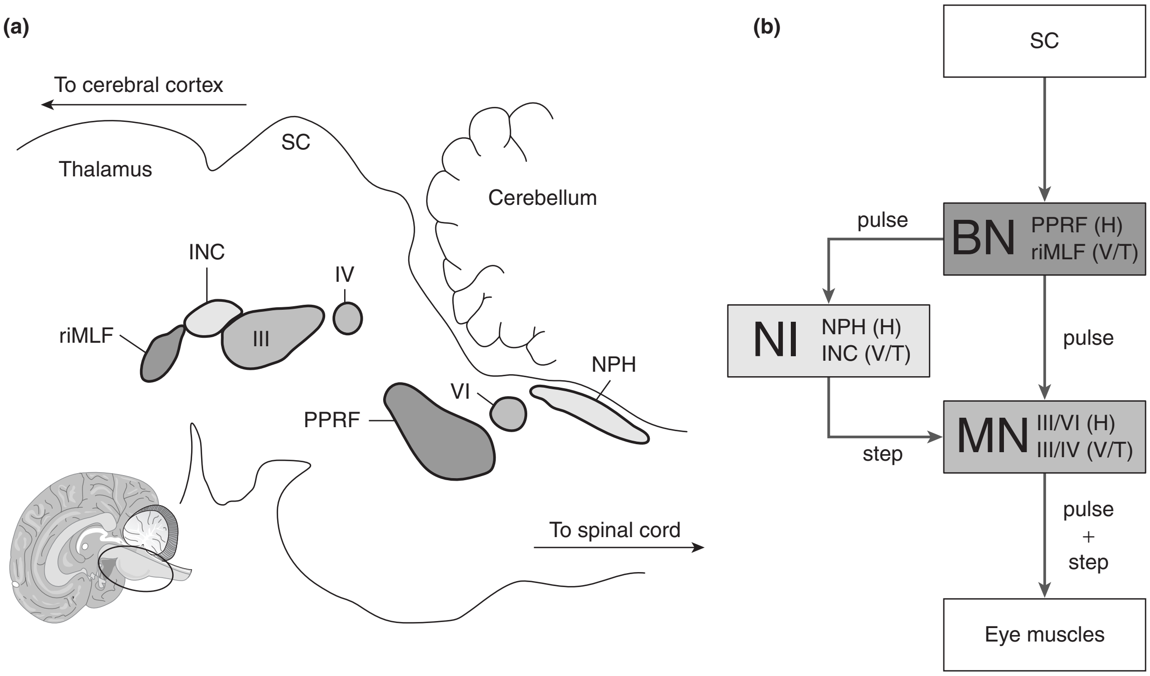


Figure 1 Brain Stem Pathways for Gaze Control

Notes: (a) Midsagital view of the primate brain stem. Burst neurons are shown in dark gray, neural integrators in light gray, and motoneurons in medium gray. The superior colliculus (SC), thalamus, and cerebellum are also shown. Figure modified from Henn, V., Buttner-Ennever J.A., & Hepp, K. (1982). The primate oculomotor system. I. Motoneurons. A synthesis of anatomical, physiological, and clinical data. *Human Neurobiology, 1,* 77–85. Inset shows the location of the brain stem. (b) A schematic of gaze shift commands traveling through the brain stem. Color scheme same as in (a). Horizontal components are indicated by (H), while vertical/torsional components are indicated by (V/T). SC = superior colliculus, BN = burst neurons, NI = neural integrator, MN = motoneurons, III = oculomotor nucleus, IV = trochlear nucleus, VI = abducens nucleus. See also color insert, Figure 33.

components. Also, within each SC, upward movements are represented in more medial sites, whereas downward movements are encoded in more lateral SC sites. These well-defined, topographical maps of space provide a representation of the spatial location of a target relative to gaze.

Additional Brain Stem Premotor Circuitry for Gaze Shifts

The brain stem also possesses a class of cells called *omnipause neurons* that fire during fixations but cease firing during saccades in any direction. These neurons are thought to dampen burst neuron activity to prevent changes in gaze during fixations and may in turn be under the control of the SC fixation zone. Together with the neural integrator, these cells make up a fixation system for holding eye position in place.

Besides brain stem neurons, several other subcortical structures contribute to gaze saccades. The thalamus relays gaze-related signals to the cortex. Inhibitory inputs from the basal ganglia (a group of structures between the brain stem and the cortex) to the SC are selectively reduced/released during saccades to help initiate the correct eye movement. Finally, the cerebellum plays a role in the upkeep of proper saccade signals. It uses sensory feedback from one's own movements to adapt the size of future saccades to be bigger or smaller as required. This ensures that (despite aging or weakening of the eye muscles) saccades can reach and stay on the right target. The cerebellum actually does this in two ways. First, parts called the dorsal vermis and fastigial nuclei ensure that the amplitude of the pulse is correct so that the eye gets to the right target. Second, the part called the flocculus makes sure that the step is properly matched so that final eye position is stable.

Other Types of Eye Movement

Smooth pursuit movements change the direction of the gaze when a target is continuously moving in space (e.g., while watching a pendulum swing). The movement of the target itself is detected by cortical neurons (e.g., medial temporal and medial superior temporal areas), and their output is sent to nuclei in the brain stem. From here, they influence the cerebellar flocculus, where neurons encode the velocity signal necessary to keep the eyes moving along with the target. Then, this velocity signal is transmitted to the same brain stem areas responsible for saccadic gaze shifts, including the burst neurons (pulse) and the neural integrators (step).

As previously described, the VOR does not change gaze direction but rather serves to keep the image of an object on the fovea when the head moves. Vestibular signals about head motion from the inner ear project directly to the motoneurons to create the pulse and indirectly to the neural integrator to produce the step that moves the eyes relative to the head. The optokinetic reflex (OKR) is similar to the VOR in that it allows the eyes to remain on target. But while the VOR uses head motion inputs, the OKR works when the head is stationary by using visual inputs when the observed scene is moving (e.g., tracking telephone poles while traveling in a car). Finally, when you tilt your ear toward your shoulder, *ocular counteroll* causes the eyes to tilt slightly in the opposite direction.

Whereas saccadic gaze shifts, pursuit movements, VOR, and OKR all move both eyes in the same direction (i.e., version movements), vergence movements move the eyes in opposite directions in order to shift gaze onto objects that are closer or further away from where one is currently looking (e.g., one verges the eyes when trying to look at one's nose). These vergence movements are controlled by the Edinger-Westfall nucleus located in the brain stem.

Higher-Level Uses of Subcortical Visual Signals

When looking around, one perceives that the world is stable (in contrast to the unstable scene produced by a hand-held video camera). This perception is possible because we are able to take into account any shifts in gaze that occur between seeing an object and acting on it. This process is called spatial updating because most investigators think that spatial information in the brain (for perception, eye movements, reach plans, etc.) is updated every time we move our eyes. Subcortical structures are thought to contribute to this by sending a copy of the gaze movement command (including its amplitude and direction) up to higher cortical centers. Proprioceptive feedback from the eye muscles may also provide information about current eye position. Similar signals may also be used in transsaccadic integration, in which captured images from several individual fixations are merged into one coherent worldview.

John Douglas Crawford and Eliana M. Klier

See also Eye Movements: Physiological

Further Readings

Buttner-Ennever, J. A., & Butler, U. (1988). Neuroanatomy of the oculomotor system: The reticular formation. *Reviews of Oculomotor Research, 2,* 119–176.

Crawford, J. D., Martinez-Trujillo, J. C., & Klier, E. M. (2003). Neural control of three-dimensional eye and head movements. *Current Opinion in Neurobiology, 13,* 655–662.

Hepp, K., Vilis, T., & Henn, V. (1988). On the generation of rapid eye movements in three dimensions. In B. Cohen & V. Henn (Eds.), *Representation of three-dimensional space in the vestibular, oculomotor and visual systems* (pp. 140–153). New York: Annals of the New York Academy of Sciences.

Klier, E. M., & Angelaki, D. E. (2008). Spatial updating and the mechanism of spatial constancy. *Neuroscience, 156*(4), 801–818.

Leigh, J. R., & Zee, D. S. (Eds.). (1999). *The neurology of eye movements.* New York: Oxford University Press.

Robinson, D. A. (1989). Integrating with neurons. *Annual Review of Neuroscience, 12,* 33–45.

Tilikete, C., & Pélison, D. (2008). Ocular motor syndromes of the brain stem and cerebellum. *Current Opinion in Neurology, 21,* 22–28.

Wurtz, R. H., & Albano, J. E. (1980). Visual-motor function of the primate superior colliculus. *Annual Reviews of Neuroscience, 3,* 189–226.

VISUAL PROSTHESES

See Prostheses: Visual

VISUAL RECEPTORS AND TRANSDUCTION

Visual perception provides us with a wealth of information about the world around us. Eyesight enables us to perform essential activities, such as walking, reading, or driving a car, and its loss can have devastating effects on quality of life. Light enters our eyes and is focused by a lens at the back of the eye onto a thin layer of highly organized neurons forming the retina. The retina is the only visible part of our brain and is the organ that allows us to see. Within the retina, incident light is detected and transformed into an electric signal by specialized neurons called photoreceptors. This signal is then processed as it travels through the layers of retinal neurons and is eventually transmitted by the axons of the retinal ganglion cells to the brain to produce our visual perception. Image-forming vision is initiated in the retina by the activation of two types of photoreceptors, rods and cones. These photoreceptors allow us to perceive the details of an image projected onto the retina and detect characteristics such as shapes, colors, and motion. Non-image-forming vision is mediated by a recently discovered type of visual receptor consisting of a subset of intrinsically photosensitive retinal ganglion cells. Signals from these receptors are involved in the contraction of the pupil in bright light (pupil reflex) and in adjusting our body rhythms (circadian clock) to day and night. This entry covers the properties, structure, and function of rod and cone photoreceptors, and photoreceptor signaling.

Properties of Rod and Cone Photoreceptors

Rods and cones have complementary distribution in most species. In humans, cone spatial density is highest in the central area of the retina, known as the fovea. The human fovea is specialized for sharp vision. Indeed, whereas the fovea occupies only 2% of the total retinal area, it contains about 1/3 of all ganglion cells. As a result, our central vision provides us with more detail than our peripheral vision. A simple example of the different resolutions of our central and peripheral vision is our ability to count the number of fingers on our stretched hand when it is directly in front of our eyes but not when it is extended to our side. Cone density rapidly declines away from the center of the fovea and is low in the periphery of the retina. In contrast, the spatial density of rods in the center of the fovea is zero, but it is high in the periphery of the retina.

In addition to their complementary distribution, rods and cones exhibit complementary functional properties that expand the overall range of our visual perception. Rods are extremely sensitive to light. Using psychophysical experiments, Selig Hecht was able to demonstrate some 70 years ago that that the simultaneous activation of 7 rods was sufficient for light perception and concluded that a rod could be activated by a single photon. After the development of techniques for recording the electrical responses from individual photoreceptors in the 1970s, single photon responses from rods were indeed observed. Their high sensitivity makes rods perfectly suited for dim light conditions, such as between dusk and dawn, and our dim light vision is mediated exclusively by the rods. The high sensitivity of our rod vision is further enhanced by the convergence of their signals. Thus, between 15 and 50 rods send their signals to a single bipolar cell. The signals from multiple bipolar cells are then combined so that one ganglion cell sums the signals from hundreds of rods before sending it to the brain. The trade off of this convergence is the low spatial resolution of our rod vision.

The high amplification of the rod phototransduction cascade that allows rods to achieve high sensitivity and to function so well in dim light conditions also results in their saturation in brighter conditions, making rods unable to respond to further increments in light. As a result, rods do not contribute to our vision during most of the day. In addition, rods experience a long refractory period following exposure to bright light and can take up to an hour in darkness to completely recover their sensitivity. This process, known as dark adaptation, can be experienced when entering from a brightly lit environment into a dark room, such as a movie theater. As our rods are initially desensitized and unable to respond to light following exposure to bright light, we are momentarily blinded and only regain the ability to see in the dimly lit theater as our rods gradually adapt to the darkness over several minutes.

Individual cones, on the other hand, are up to 100-fold less sensitive than rods and are, therefore, useless in dim light conditions. Their lower sensitivity, however, makes cones perfectly suited for brighter light conditions, such as during the day. In addition, the use of three types of cones with different spectral sensitivities allows our visual system to discriminate colors and gives us color perception in bright light. The inability of cones to detect dim light is what deprives us from color vision at night. In contrast to the signal integration from hundreds of rods, cone signal integration is minimal and signals from individual foveal cones can be relayed to the brain. As a result, the spatial resolution of our central vision, driven primarily by the cones, is excellent.

Rod and cone photoresponses have different kinetics that result in complementary temporal resolution of our rod and cone vision. Whereas mammalian rod responses peak in about 100 milliseconds (ms) and last about 1 second, cone responses are typically several times faster. The rapid activation and subsequent inactivation of cone responses provides the basis for the high temporal resolution of our cone-mediated vision. In contrast, the slower rod responses limit the temporal resolution of rod-mediated vision but enable rods to sum signals over longer periods of time, enhancing sensitivity in dim light conditions.

Unlike rods, cones dark adapt rapidly and can fully recover their sensitivity within a few minutes following exposure to bright light. In addition, cones have a remarkable ability to adjust their sensitivity and remain photosensitive, even in extremely bright light. This process, known as light adaptation, prevents cones from saturating in bright light and allows us to see in the extremely wide range of light intensities that we encounter throughout the day. With rods saturated, cones are responsible for most of the visual information reaching our brain during the day. This is why cone disorders, such as macular degeneration, the most common cause of blindness in the elderly, have such devastating effects on vision.

Structure and Function of Rod and Cone Photoreceptors

Rods and cones have similar structures; they both contain an elongated outer segment (where the conversion of light into an electrical signal takes place), an inner segment, a cell body, and a synaptic terminal. The outer segment is a specialized ciliary compartment that consists of hundreds of membrane discs stacked like a roll of coins perpendicular to the longitudinal axis of the outer segment. These discs contain the visual pigment, the light-sensitive molecule of photoreceptors that triggers the activation of the phototransduction cascade (see next section) upon the absorption of a photon. The visual pigment is packed in the disks of the outer segment at a very high concentration, 3.5 millimolars (mM), which results in a ~40% probability of a photon being absorbed as it travels through the outer segment of a single photoreceptor. In rods, the disks are surrounded by, but separate from, the plasma membrane. In cones, these disks are formed from foldings of the plasma membrane. The resulting higher membrane surface area of cone outer segments is believed to be important for the rapid flow of molecules in and out of cones. The outer segments undergo continuous renewal as new discs are formed daily at their base, while discs at the tip are shed and digested by the adjacent pigment epithelium. The lifetime of a disk, from its formation to its shedding in mammalian rods is about 10 days. Disk shedding is believed to play a role in the removal of toxic photooxidation products and the turnover of phototransduction proteins.

At the other end of the cell, the synaptic terminal connects rods and cones to secondary neurons in the first step of relaying the signal from photoreceptors to the brain. The absorption of a photon by a molecule of visual pigment in the outer segment triggers the phototransduction cascade and ultimately results in change in membrane potential, the voltage across the plasma membrane of the photoreceptor. This, in turn, modulates the release of neurotransmitter from the synaptic terminal of the cell. In darkness, rods and cones are depolarized to about –35 millivolts (mV) and, as a result, their synaptic terminals continuously release neurotransmitter, glutamate, onto the synapse of the second order bipolar cells. In most neurons, activation produces depolarization and release of neurotransmitter at their synaptic terminal. In contrast, light activation of vertebrate rods and cones results in their hyperpolarization and the reduction or complete block of neurotransmitter release from their synaptic terminals. Thus, activation of photoreceptors by light is reported to the second order neurons as reduction in the release of neurotransmitter from their synapses.

Photoreceptor Signaling

In rod and cone photoreceptors, light detection is achieved by a series of reactions, called phototransduction cascade (Figure 1), which converts the energy of light into a change in the membrane potential of the cell. Photoreceptor activation is initiated when a photon is absorbed by a molecule of visual pigment. In both rods and cones, the visual pigment consists of protein, opsin, covalently attached to a light-sensitive molecule called chromophore, a derivative of Vitamin A. Similarly to olfactory receptors, the opsin molecule spans the disk membrane seven times, forming a barrel-like structure. A pocket in the core the protein contains the chromophore, which is covalently attached to opsin. Absorption of a photon by the chromophore, which is bent, triggers its straightening. This change in the shape of the chromophore in turn induces rearrangement of opsin to activate the pigment molecule. Once activated, the visual pigment binds to and activates a heterotrimeric G protein called transducin (Figure 1). As one visual pigment molecule can activate dozens of transducin molecules during its active state lifetime, this represents the first step of phototransduction amplification. The next step in the phototransduction cascade involves the activation by transducin of an effector enzyme, cGMP phosphodiesterase (PDE), and the hydrolysis of the phototransduction cascade second messenger molecule, cGMP. In darkness, cGMP binds to and keeps open cGMP-gated ion channels. These transduction channels allow the influx of positively charged sodium and calcium ions and keep photoreceptors depolarized. The light-induced decline in cytoplasmic cGMP concentration due to hydrolysis by PDE leads to the closure of the cGMP-gated channels in the outer segment (Figure 1). This blocks the influx of positive ions into the outer segment and makes the photoreceptor membrane potential more negative, so that the cell becomes hyperpolarized. The hydrolysis of multiple cGMP molecules by a single PDE represents a second amplification step in the phototransduction cascade. Response termination and resetting of the photoreceptor for subsequent photon detection is achieved as the visual pigment, transducin, and

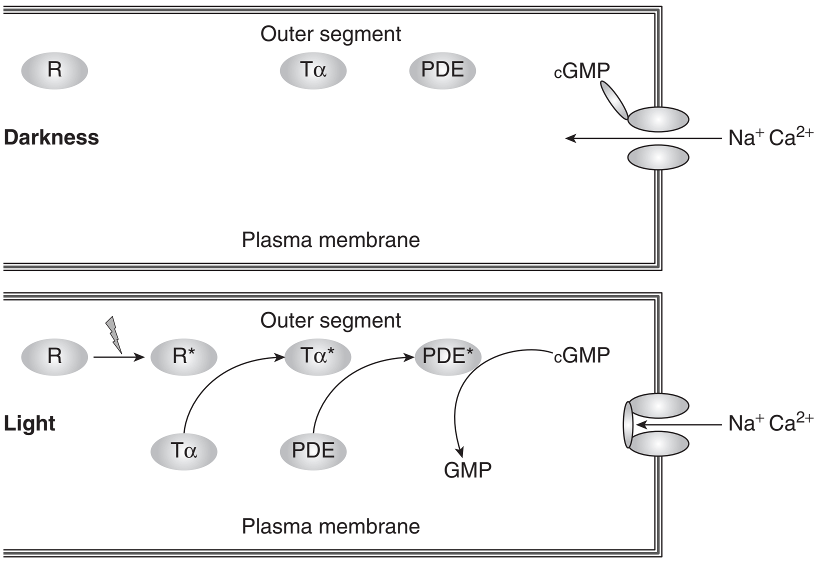


Figure 1 Phototransduction in Photoreceptors

Notes: Photoreceptors rely on a cascade of reactions to convert the absorption of a photon by a visual pigment molecule to the closing of ion channels in the plasma membrane for the generation of an electrical signal. This figure shows a simplified scheme of phototransduction cascade activation in outer segments of photoreceptors. R = rod visual pigment, also known as rhodopsin; Tα = transducin, PDE = phosphodiesterase. The active state of each protein is indicated by (*).

PDE are inactivated and the concentration of cGMP is restored to its dark level by guanylyl cyclase (GC).

The activation of the receptor molecule, the visual pigment, generates an electric signal by a chain of reactions. This use of a cascade of reactions offers several advantages. First, the cumulative amplification of the signal at several consecutive phototransduction steps allows rods to achieve their high sensitivity. Next, a second messenger mechanism allows pigment activation in the lumen of the outer segment to be relayed to the channels in the plasma membrane, thus vastly increasing the detecting area of photoreceptors. Finally, the ability to modulate multiple phototransduction steps also allows for powerful regulation of sensitivity, or adaptation, in various light conditions. This adaptation is crucial for our ability to perceive light ranging over 14 orders of magnitude because the range of light sensitivity in dark-adapted photoreceptors is rather narrow at about 100-fold, or 2 orders of magnitude.

Vladimir J. Kefalov

See also Color Perception: Physiological; Eye: Structure and Optics; Vision; Visual Light- and Dark-Adaptation; Visual Processing: Retinal

Further Readings

Arshavsky, V. Y., Lamb, T. D., & Pugh, E. N., Jr. (2002). G proteins and phototransduction. *Annual Review of Physiology, 64,* 153–187.

Burns, M. E., & Baylor, D. A. (2001). Activation, deactivation, and adaptation in vertebrate photoreceptor cells. *Annual Review of Neuroscience, 24,* 779–805.

Ebrey, T., & Koutalos, Y. (2001). Vertebrate photoreceptors. *Progress in Retinal and Eye Research, 20,* 49–94.

Fain, G. L., Matthews, H. R., Cornwall, M. C., & Koutalos, Y. (2001). Adaptation in vertebrate photoreceptors. *Physiological Reviews, 81,* 117–151.

Fu, Y., & Yau, K. W. (2007). Phototransduction in mouse rods and cones. *Pflügers Archive: European Journal of Physiology, 454,* 805–819.

Pugh, E. N., Jr., & Lamb, T. D. (1993). Amplification and kinetics of the activation steps in phototransduction. *Biochimica et biophysica acta, 1141,* 111–149.

Saari, J. C. (2000). Biochemistry of visual pigment regeneration: The Friedenwald lecture. *Investigative Ophthalmology & Visual Science, 41,* 337–348.

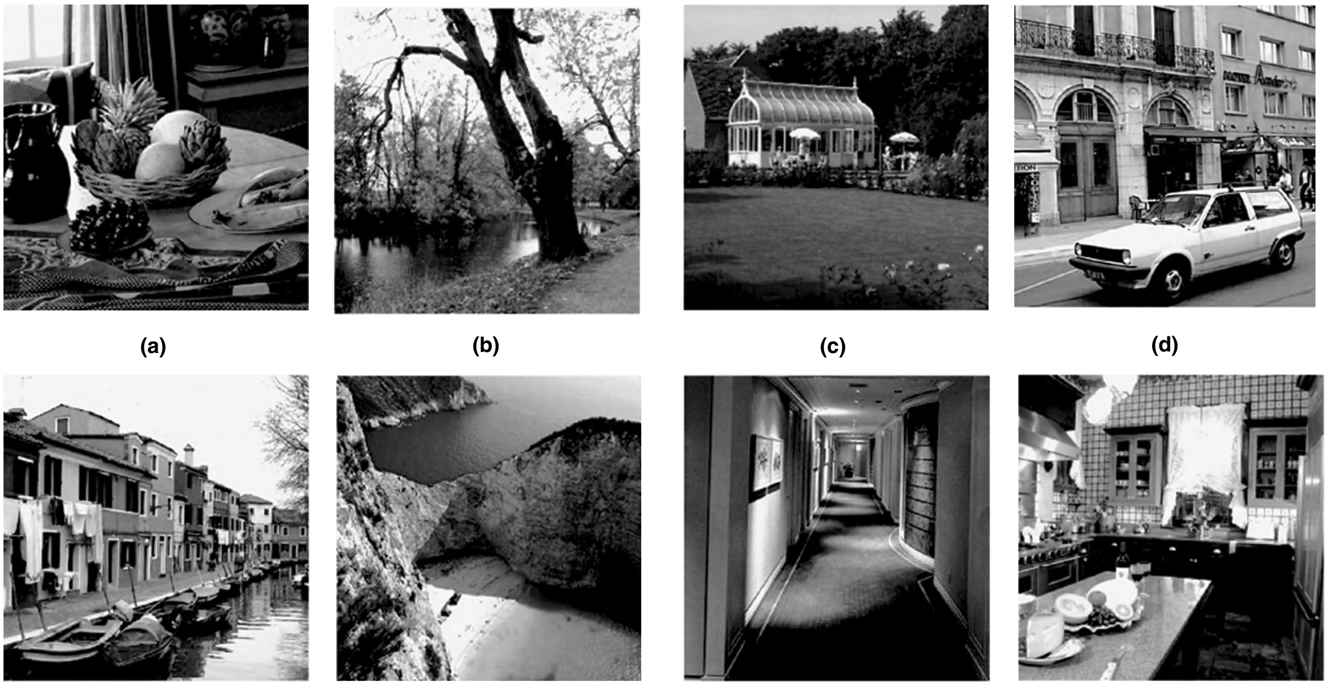
VISUAL SCANNING

See Eye Movements: Behavioral

VISUAL SCENE PERCEPTION

A visual scene is commonly defined as a view of an environment comprised of objects and surfaces organized in a meaningful way, like a kitchen, a street, or a forest path. More broadly, the domain of scene perception encompasses any visual stimulus that contains multiple elements arranged in a *spatial layout,* for example, a shelf of books, an office desk, or leaves on the ground. As a rough distinction, objects are typically acted upon, whereas scenes are acted within. Most visual elements in the world will be categorized as either an object or scene, although an item's status may change depending on how it is used: A keyboard is readily regarded as an object when being purchased, for example, but when placed on the desk, it becomes part of a continuous surface and layout, and a scene (for your hands), when you are trying to find the right key.

The complex arrangement of objects and surfaces in natural scenes can create the impression that there is too much to see at once. However, we are able to interpret the meaning of multifaceted and complex scene images—a wedding, a birthday party, or a stadium crowd—in a fraction of a second. This is about the same time it takes a person to identify that a single object is a face, a dog, or a car. This remarkable feat of the human brain can be experienced (and enjoyed) at the movies: With a few rapid scene cuts from a movie trailer, it seems as if we have seen and understood much more of the story in a few instants than could be described later. We will easily remember the movie's genre and limited context (for example, a romantic story with views of Venice or a science fiction story set



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 1 Look for one second only at each picture; close your eyes for a few seconds and then view Figure 2

in the near future), but we will have forgotten detailed information. The same phenomenon happens when quickly changing television channels or flipping pages of a magazine: One single glance is often enough to recognize a popular TV personality, a high-speed car chase, or a football game, but memory of the details is wiped away almost immediately. Perceiving scenes in a glance is like looking at an abstract painting of a landscape and recognizing that a "forest" is depicted without seeing necessarily the "trees" that create it. This entry discusses perceiving and remembering visual scenes and mechanisms of visual scene perception.

Perceiving and Remembering Visual Scenes

Because a scene can encompass a large space, we must acquire information about its extent by navigating our bodies and moving our head and our eyes. Although we seem to experience a continuous world, the brain actually samples the visual world in a series of "snapshots," by moving the eyes three times a second. Perceiving a visual scene is like watching a movie in which the camera cuts quickly from one view to the next. During each brief moment in which we see a particular view, we are able to understand its overall meaning or "gist" (e.g., a wedding, a dog running in a park, a busy New York street), we have a compelling perception of space, and we anticipate what is coming next. However, details and objects are quickly lost from visual memory and, in many cases, are not even perceived in the first place.

To illustrate, look at Figure 1. It contains photographs of 8 scenes. Spend about one second on each image, then close your eyes for a few seconds, and then turn the page to view Figure 2. There you will find a set of 8 photographs. Your task is to determine which images you saw in the first figure. Come back to reading this section after doing the test.

Some of the images in Figure 2 look remarkably similar to images from Figure 1, but they are, in fact, all different. However, most people misidentify some or all of the images in Figure 2 as being the same images from Figure 1. Here, we use the limitations of visual memory to reveal a number of important properties of scene perception. Because there is so much information to perceive at once, we link the visual input with our memories and expectations, leave out some details, and transform what we perceive.

The Boundary Extension Phenomenon

The pair of images that show a set of objects on a table (Figures 1a and 2a) illustrates a systematic error in scene perception revealed by Helene Intraub's work: People often remember a wider view of a scene than was originally seen. If you look closely at the borders of the images, you will notice that Figure 2(a) is a wider view in which you can see more of the pitcher and the tablecloth. This boundary extension effect occurs as early as half a second after viewing an image, and may also occur between two views of the same place. For instance, the autumn park images (Figures 1b and 2b) depict the same place, but in the second image, the observer has moved a few steps closer to the riverbank. In many cases, people will not notice the difference between two views of the same place that differ by 10 to 20° of head movement or a few steps of body translation.

These spatial memory errors reveal a fundamental mechanism of visual scene analysis: People rely on their previous experience and knowledge of the world to rapidly process the vast amount of detail in a real world scene. One's current view of a scene is automatically incorporated into a "scene schema," which includes stored memories of similar places that have been viewed in the past, as well as expectations about what is likely to be seen next. Although we aren't aware of it, viewing a scene is an active process in which perceived images are combined with stored knowledge to create an internal reconstruction of the visual world.

The Change Blindness Phenomenon

The pictures of a house and lawn (Figures 1c and 2f) and the pictures of a city street (Figures 1d and 2d) illustrate another well-known phenomenon of visual scene perception: People are surprisingly poor at detecting a change in a scene when the change happens between two eye movements or a shift in viewpoint (like a turn of the head or, as in this demonstration, a turn of a page). This demonstration illustrates the change blindness phenomenon studied by Daniel Simons and Ronald Rensink. In Figure 2, the scene depicting a house and lawn is missing the parasols and people in the center, and the two images of the European street differ actually on half a dozen details (the whole building on the right, the presence of pedestrians, and more). The change blindness phenomenon further illustrates that, contrary to our subjective experience, many details of the visual world simply go unnoticed.

The Gist Phenomenon

If you thought that the canal streets (Figures 1e and 2g), the mountainous coasts (Figures 1f and 2e), or the corridors (Figures 1g and 2h) were the same images, then you may have experienced an error based on scene gist. Different images that share a similar meaning and look similar may be falsely remembered as the same scene. This phenomenon is similar to the feeling of déjà vu, when a novel place is experienced as familiar. In a similar vein, it should have been easy to spot that the living room of Figure 2c was a new scene, as there are no similar images, in terms of spatial layout or semantics in the set of Figure 1.

Mechanisms of Visual Scene Perception

Research in scene perception has traditionally treated objects as the atoms of recognition. However, studies have shown that the speed and accuracy of scene recognition are not affected by the number of objects in the scene, which suggests an alternative: The meaning of scenes may be quickly inferred from their spatial layout, the arrangement of surfaces and forms.

The Prevalence of Global Spatial Layout

Color insert Figure 34(a) illustrates how the spatial arrangement of surfaces and regions drives scene perception. When looking at the blurred image on the left, viewers describe the scene as a car parked in front of a fancy building's facade. Even though the local information available in the image is insufficient to recognize the objects individually (the car, street, and facade are too blurry to be recognized in isolation as illustrated in the center of color insert Figure 34a), viewers are confident and highly consistent in their descriptions. This is because the blurred scene has the spatial layout of a typical street.

Figure 2 Which pictures did you see in Figure 1?

When the image is shown in high resolution, new details reveal that the image has been manipulated and that the buildings are in fact pieces of furniture. In fact, more than half of the image depicts an indoor scene, not a street. The misinterpretation of the low-resolution image is not a failure of the visual system. Instead, it illustrates the strength of spatial layout information in determining the identity of the objects within a scene. The importance of global layout information is especially evident in degraded viewing conditions (for example, viewing a scene at a distance or at a quick glance) in which object identities cannot be derived from local information alone.

Segmentation and Figure-Ground Analysis of Visual Scene

Color insert Figure 34(b) demonstrates further the dominance of global information in scene perception. The left image seems to show a forest scene, with the observer looking up through the trees to see the sky in the background. But if you study the image closely, you may notice local inconsistencies, such as incorrect shadows and odd surface shapes or occlusions. The right image of color insert Figure 34(b) is the same image, but upside down: now the forest scene depicts a river receding into the distance. In this particular illusion, reversing the image changes the assignation of elements as *figure* (i.e., the object of focus) or *ground* (i.e., the rest of the perceptual field). When the forest scene is turned upside down, the sky becomes a river and the darker foliage in the distance (a figure element in the original view), turns into the reflection of trees in the water (a ground element in the upside-down view). The attribution of scene elements as figure or ground depending on the perceived layout shows that perceiving a complex, real world scene is an interaction between bottom-up and top-down processing. The visual features extracted from the image (e.g., color, lines, and patterns of texture) are rearranged according to our expectations and knowledge of the world. For instance, we know that a blue region at the top of an outdoor scene is most likely the sky; and we know that, in a large scene, elements below the horizon line are more likely to be close-by than elements above the horizon line.

The Role of Statistical Regularities in Visual Scene Perception

The visual world is not random: objects, surfaces, and the gaps between them are organized in predictable fashion. Just as there are rules governing the structure of objects (for example, faces have two eyes, a nose, and a mouth), the structure of many real-world scenes is governed by strong configural rules that predict what kind of objects and surfaces will appear near each other. This is illustrated in color insert Figure 34(c). By averaging hundreds of images aligned on either a sailboat (left) or a cow (right), a common pattern of intensities emerges around each object. The average sailboat is surrounded by a blue background (water) with lighter patches at the top and sides (sky or shore) and vertical elements that hint at neighboring boats. The average cow, on the other hand, is surrounded by greenish patterns representing fields. This averaging illustrates the statistical regularities that are imposed on the scene by the objects within it.

Similarly, environments that share the same category and function tend to share similar layouts: a highway is a flat ground surface stretching to the horizon, affording speedy travel; a corridor is an enclosed, perspective space, with an unobstructed path to afford navigation; cities are made up of vertical facades; and a dining room is a space organized around a central object, a table. Because we experience scenes as real, three dimensional spaces, Aude Oliva and Antonio Torralba have proposed that visual scene perception is based on a global layout representation that describes the space and volume of the scene (e.g., beaches are large, open environments; closets are small, enclosed spaces) and not necessarily the objects the scene contains.

Importantly, human observers are very good at extracting summary statistics of scenes that help with recognition of the space and the objects within: that is, they are able to rapidly estimate the regularities (and differences) in the layout of a scene or in the objects that compose it. For example, people know the mean size of a collection of objects without knowing the precise size of each of them; they know the center of mass of a group of objects without necessarily remembering the location of each of them; and they can rapidly recognize whether a space is a small or a large volume, or a busy or an empty space before identifying whether it is a street or a closet. Summary statistics and other regularities of the environment learned with experience are valuable for scene perception: They provide an efficient and compact representation of the image that can give information about other scene properties and facilitate object perception and search before local image information has been fully analyzed. This is experienced every day: A glance at your own kitchen or bedroom cues the location of objects that may not be immediately visible. For example, a glimpse of your bedroom should be sufficient to cue you to the location of the alarm clock in your mental representation of that place, even if the alarm clock isn't visible from your current position.

Conclusion

How do we perceive visual scenes? Decades of behavioral research suggest that scene perception begins at a global level. First, the spatial layout and observer's viewpoint are evaluated, and then the localization and recognition of parts and objects within the scene progress at a slower rate. In other words, you would know that you are in a large kitchen, before recognizing that a particular form is a fridge and that a rectangular object is a microwave. Scene perception is also by nature a constructive process that links visual percept to stored representation: It is often prone to errors and distortions, but these false reconstructions reveal how the brain analyzes complex visual information. Environments have all sorts of regularities that we learn and store in memory. When faced with a novel scene, we use our knowledge and expectations to rapidly understand its meaning, although this comes at the cost of losing some detail from perception. Scene perception is essentially the process of reconstructing a space from a lifetime of stored representation.

Aude Oliva

See also Auditory Scene Analysis; Change Detection; Computer Vision; Eye Movements: Behavioral; Navigation Through Spatial Layout; Perceptual Organization: Vision; Perceptual Segregation; Rapid Serial Visual Presentation; Spatial Layout Perception, Neural; Spatial Layout Perception, Psychophysical; Statistical Learning; Top-Down and Bottom-Up Processing; Visual Memory; Visual Scene Statistics; Visual Spatial Frequency Analysis

Further Readings

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14,* 143–177.

Chun, M. M. (2003). Scene perception and memory. In D. Irwin & B. Ross (Eds.), *Psychology of learning and motivation: Advances in research and theory: Cognitive vision,* (Vol. 42, pp. 79–108). San Diego, CA: Academic Press.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392,* 598–601.

Greene, M. R., & Oliva, A. (2009). Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology, 58,* 137–179.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50,* 243–271.

Intraub, H. (2007). Scene perception: The world through a window. In M. A. Peterson, B. Gillam, & H. A. Sedgwick (Eds.), *The mind's eye: Julian Hochberg on the perception of pictures, films, and the world* (pp. 454–466). New York: Oxford University Press.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*(3), 353–383.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Science, 11*(12), 520–527.

Potter, M. C. (1999). Understanding sentences and scenes: The role of conceptual short-term memory. In V. Coltheart (Ed.), *Fleeting memories: Cognition of brief visual stimulus* (pp. 13–46), Cambridge: MIT Press.

Schyns, P.G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science, 5,* 195–200.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9*(1), 16–20.

VISUAL SCENE STATISTICS

Visual scene statistics refers to the study of the regular patterns that recur in visual images and the random fluctuations that perturb these patterns. For the purpose of this entry, "statistics" simply refers to a way of describing a phenomenon that accounts for both those aspects that we can be sure about as well as those aspects that are beyond our ken. This is not a new idea. As long as people have left written records they have appreciated the fact that the world around us contains things that are generally predictable but at the same time capable of surprising us. The sun always rises, but a given day is sunny or not according to its whim.

We have long known that major features of perceptual systems adapt to major features of the environment—birds that live in clear air can see very well but fish that live in murky waters have good vibratory and electrical sensation—but it turns out that many more subtle features of visual scenes have also influenced both the evolution and development of our visual systems, and most of these features are not completely predictable, but rather occur as long-term statistical trends across a large number of visual scenes. These visual scenes, which of course provide the input for our visual systems, thus require a *statistical* description because they have both these substantial regularities as well as large components of random variation.

To begin with an almost trivial example, the most fundamental visual scene statistic is arguably the fact that any given spot on earth gets brighter and darker as the earth rotates. Thus, nocturnal animals have a relatively large number of very sensitive rod photoreceptors, but relatively few of the cone photoreceptors on which we humans predominantly rely. However, there is also a large random component to how bright the world is at any given time, mostly due to the weather and other atmospheric factors (e.g., forest fires or volcanic eruptions). Thus, crepuscular species (active during dawn and dusk—cats, for example) need a visual system that works well at the light level of a typical dawn, but that can also tolerate fairly large fluctuations; they need to function during exceptionally bright and dark mornings, even if that means sacrificing some performance on the more typical days. Similarly, diurnal animals (humans, for example) need a visual system that can function not only on typical days, but also during exceptionally dark cloudy days at high latitudes. The point is that, because visual scenes contain random fluctuations, visual systems must be tuned to cope with not only average conditions, but also with the inevitable fluctuations about those average conditions.

There are a few other obvious scene statistics that have a powerful impact and may be familiar to most of us. One is that light tends to come from above. It is no accident that children enjoy making their faces look scary by lighting them from below with a flashlight. A face lit from below should be scary because it something that almost never occurs naturally. The visual system's assumption of light coming from above leads to some powerful illusions—one that is easy to make in any drawing or presentation program (such as PowerPoint or Keynote) is the following. Draw a middle-gray background, then draw a nickel or quarter sized circle on the background, and shade the circle with a gradient that runs from white on top to black on the bottom. The circle will appear to pop out from the background (particularly if you step back from your computer a little ways and close one eye), as though somebody had pushed out an indentation from the back of the screen with his or her thumb. Because light generally comes from above, something convex should be lighter on top and darker (in shadow) on the bottom. Our visual system, acting as a good statistician should, builds on this assumption to make sense of the world, even though the assumption isn't always correct.

More recently, we have learned that many additional properties of visual scenes have important statistical regularities despite fairly large scene-to-scene variation, and that these regularities likely have profound implications for visual perception. The most basic of these is perhaps the spectrum of wavelengths of light (which we perceive as different colors) to which visual systems of different species, and quite possibly different groups of humans, are typically exposed. This results in different distributions of photopigments, the molecules in the light-sensitive receptors in the back of the eye, in different species inhabiting differing ecological niches.

Important recent discoveries about visual scene statistics are also a bit subtler than the previous examples and involve what are called "conditional probabilities." Consider taking a sample of the visual world using a video camera (i.e., a series of video frames in rapid succession). Now take an arbitrary pixel location in a few successive frames. Given the world in which we live, what can we say about the value of this pixel from picture to picture or in relation to other pixels in the same frame? Certainly, if we look at any given picture, we see a rich array of objects, but that is because we use higher centers of visual processing to look at the picture as a whole and imbue parts of it with meaning. But if we were forced to zoom in and look at the picture on a pixel-by-pixel basis, it is not clear that we could glean much information from it at all. It turns out, however, that there is a rich statistical relationship between any given pixel and its neighbors, both in space (nearby pixels in one frame) and in time (the same pixel in nearby frames). If this were not true—if the color of different locations is space and time was unrelated—our entire world would look like the "snow" on a television with no reception. Our world, however, is made of objects, and pieces of any given object tend to reflect light in a similar manner, whereas pieces of different objects tend to reflect light differently. Thus, if a given pixel is light green, it's a good bet—literally—that nearby locations will also be light green (perhaps because they all belong to the same leaf), but for more distant pixels the bet becomes less certain, corresponding to the increasing probability that the pixel belongs to trunk of the tree, or a different kind of leaf. This statistical property of visual scenes allows the visual system to save a lot of work by making these types of bets. If a given location is light green, why waste valuable brainpower determining that an adjacent location is light green when that can just be assumed with a high probability of being correct? The brain does, in fact, do just that. Neurons in the primary visual cortex (the cortical receiving station for visual input from the eyes) are perfectly tuned—through some combination of evolution and development—to make the right bets, on average, given the physics of the world in which we live and the way this world produces images in our eyes.

Not only do visual scene statistics allow our visual system to make bets on the colors and brightness of neighboring locations in a scene, but they also allow it to make bets on the location and orientation of contours and edges in a scene. The logic is essentially the same. Our world is largely grouped into distinct objects, so these objects must have boundaries (by definition). Thus, if there is a visible contour at a given location, it's a good bet that nearby locations along the direction of the contour will also have a contour running in that direction. Moreover, if a given location has a contour running a given direction, and a more distant location also has a contour running the same direction, then it is a good bet that there is also a connecting contour, even if that contour is being hidden by something else. Surprisingly, at least to anyone familiar with the history of the psychology of visual perception, the quantitative statistical expression of the probabilities associated with these "bets" has led to a deeper, quantitative understanding of some long-standing but purely descriptive ideas, such as the Gestalt laws of perception.

More recently, visual scene statistics has shed some light on the role of stereopsis—our ability to perceive depth by virtue of our having two eyes in the front of our head—in everyday outdoor viewing. It has long been assumed, and taught in many introductory courses, that stereopsis works primarily at short viewing distances (a few feet to several yards). This was based upon the fact that the stimulus for stereopsis, the slight differences in the two eyes' views of the world (termed *retinal disparities*), decreases rapidly as the viewing distance increases. However, when the retinal disparities are actually measured where people look in outdoor scenes, one finds that these scenes are an incredibly rich source of stereoscopic information. Moreover, the statistical distribution of these naturally occurring disparities is virtually identical to the sensitivity of populations of neurons encoding disparity in the visual cortex, indicating that the brain has evolved and developed hand-in-hand with the environment to encode the stimuli present—no more, no less.

More crucial is the potential for the visual system to make inferences (bets) about important properties of the world that are themselves difficult to detect directly by using their statistical relationships with properties that are more easily detectible. An example of this is the measurement of the distance to objects, which is one of the most important things our visual system does: How far away is that oncoming car? Can I jump to that ledge? Can I hit that mastodon with my spear? Yet distances cannot be measured directly by the visual system. Retinal disparities and other sources of information give signals that are, at best, only proportional to distance. For example, if one object is partially occluding another then, obviously, one object must be closer. But, crucially, the occlusion provides no information about the actual *distances* of the objects. Recent work, however, is showing that both distances and changes in distances have subtle but reliable statistical relationships with scene properties that the visual system *can* measure directly, such as the color change across object boundaries. Thus, if the visual system detects a certain amount of color change and a certain amount of brightness change in a given small patch of a scene, it is likely that there is also a distance change of a certain amount.

Looking ahead, there are probably many other simple conditional statistics in the environment that the visual system exploits. Of course, there are probably also statistical relationships in the environment to be discovered that are more complex than the relationships between two or even three variables. Regardless, those relationships already discovered have furthered our understanding of the visual system tremendously. In summary, the evolution, development, and function of any perceptual system, including the visual system, can only be appreciated in the context of the environment in which it operates, and the study of natural scene statistics (and tasks) is absolutely necessary if we ever hope to truly understand our perceptual systems.

Lawrence K. Cormack

See also Bayesian Approach; Binocular Vision and Stereopsis; Evolutionary Approach; Gestalt Approach; Statistical Learning

Further Readings

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology, 59,* 10.1–10.26.

Graham, D. J., & Field, D. J. (2008). Variations in intensity statistics for representational and abstract art, and for art from the eastern and western hemispheres' perception. *Perception, 37*(9), 1341–1352.

Liu, Y., Bovik, A. C., & Cormack, L. K. (2008, November 19). Disparity statistics in natural scenes. *Journal of Vision, 8*(11):19, 1–14. Retrieved from http://journalofvision.org/

Yang, Z., & Purves, D. (2003). A statistical explanation of visual space. *Nature Neuroscience, 60,* 632–640.

VISUAL SEARCH

Visual search is any task where the observer has to find some desired object, or *target,* in a field of multiple nontarget *distractor* objects. We perform hundreds of visual searches each day: searching for a shirt in our closet, finding our car keys on the nightstand, or looking for a friend in a crowd of strangers. There are many socially important visual search tasks, such as searching for threats in X-ray scans of baggage, finding potential tumors in mammograms, or looking for airplane wreckage in satellite photographs. In classic laboratory search tasks, the observer sees a *search array* of items (e.g., colored rectangles) and must report whether or not a specified target (e.g., "green vertical rectangle") is present. These stimuli are intended as an easily controlled abstraction of complex scenes. Important stimulus parameters include set size (total number of items), target presence or absence, similarity between targets and distractors, and similarity among distractors.

Time and accuracy are the dependent variables in visual search tasks. Some experiments measure the time required to successfully find a target, maximizing accuracy by presenting the search array until the observer responds with a key press. Other experiments measure how accurate observers can be given limited search time. Many visual search studies also measure eye movements. The focus of the eyes (*fixation* point) is only a coarse guide to the locus of attention, and many laboratory search tasks can be carried out without moving the eyes. Nevertheless, the average number of eye movements needed to find the target is highly correlated with reaction time, and much can be learned from analyzing patterns of eye movements. This entry explores set size effects, basic features and search asymmetries, eccentricity effects, contextual effects on search performance, and major theories of visual search.

Set Size Effects

One of the most common and useful findings in the visual search literature is the set size effect: Reaction times tend to increase (and accuracy to decrease) with the number of items in the display. What is interesting here is not that the search becomes harder when there are more items to search through, but that different set size effects are observed with different stimuli. Set size effects can be measured by varying set size and obtaining the slope of the function relating reaction time to set size. This *search slope* indexes the time cost of each additional stimulus. Steeper search slopes (i.e., greater increases in response time with set size) mean less efficient searches. Shallow negative slopes can be observed in some circumstances, whereas extremely steep slopes can be obtained by making stimulus discrimination difficult and/or dependent on fixation. However, the range of search slopes for most tasks is between 0 and 150 milliseconds (ms)/item.

Slopes obtained in most laboratory search tasks fall on a continuum coarsely divided into three categories. Consider a field of vertical bars, identical except that the one stimulus (the target) is green and all of the remaining stimuli (distractors) are red. Because the target differs from the distractors by a single feature, this is termed a *feature search.* In feature searches, the search slope is zero; the target is found rapidly no matter how large the set size. In Anne Treisman's memorable phrase, the target seems to "pop-out" of the background; such searches are also called *pop-out searches.* Now imagine that the target is a green vertical bar but distractors are red vertical and green horizontal bars. Here the target is not the only green item, or the only vertical item, but the only conjunction of these features; this type of search is termed *conjunction search.* Conjunction search is less efficient than feature search, with slopes typically around 5 to 15 ms/item. Finally, imagine that each item consists of a vertical and a horizontal bar, with distractors looking like Ls (in any of the 4 canonical orientations) and the target looking like a T (similarly oriented). Here the target differs from the distractors not in any particular feature, but in the spatial arrangement of its elements. Such *spatial configuration searches* are quite inefficient, producing slopes of 25 to 50 ms/item.

Basic Features and Search Asymmetries

Search slopes can be used to infer how much information the visual system has about objects before they are attended. The more "preattentive information" is available, the more efficient a search will be. Feature search indicates that all of the information needed to find the target is available preattentively. If a visual attribute produces a flat search slope, researchers infer that neural machinery to process that attribute is available in parallel across the visual field. Such attributes are termed basic features. Well-established basic features include color, orientation, motion, and spatial frequency. Other putative basic features, such as faces or lighting direction, are currently being debated.

Search experiments also can be used to infer the preattentive representation of features. A single red item pops out in a field of green items, and vice versa. This color search is therefore symmetrical. However, a line tilted 15° from vertical pops out of a field of vertical lines, whereas a vertical line is difficult to find among tilted lines. Because it is generally easier to detect the presence of a feature than its absence, a search asymmetry of this kind suggests that, preattentively, verticality is the absence of tilt. Other search asymmetries arise because it is easier to find more among less than less among more. For example, long lines pop out of short lines but not vice versa.

There are two caveats to keep in mind when interpreting search asymmetries. First, to search for A among B may be easier than to search for B among A, even though the A among B search is not a pop-out task. In this case, the likely explanation is that it is easier to reject B distractors as "not-A" than vice versa. Thus, it is easier to find an inverted face among upright faces than vice versa, not because the inverted face has some unique feature, but because upright faces can be rejected as distractors more rapidly. Second, in order to infer an asymmetry in the visual system from asymmetric results, the experimental design must be symmetric. For example, a moving dot pops out among stationary dots, but the reverse is not true. Here, it may be only the experimental design that is asymmetric. With a moving target, stationary distractors are all the same, whereas for a stationary target, moving distractors can move in many directions.

Eccentricity Effects

Humans appear to have a default attentional distribution, which becomes more pronounced with age: All else being equal, a target further from the center of the visual field will be detected later than one closer to fixation. This is true even when search items are large enough to be easily discriminable in the visual periphery. Some of this *eccentricity effect* may be due to the fact that there is more cortical machinery devoted to processing central vision, but there also appears to be an attentional bias toward the current locus of fixation. This bias can be overcome by cueing manipulations, such as directing attention to a ring-shaped object.

Contextual Effects on Search Performance

In addition to the features of the search array, search performance is affected by events on previous trials. If the same item can serve either as a target or a distractor on different trials (*inconsistent mapping*), a search will be much more difficult than if targets never serve as distractors (*consistent mapping*). If the target appears in the same location on the current trial as on a previous trial, it will be found more quickly. The same goes for repeats of color, orientation, motion, or other features. Such *priming* effects are not limited to target features; repetition of distractor sets also improves performance. Priming can influence performance over several intervening trials.

Search performance can also be influenced by contingencies operating over many trials, a phenomenon known as *contextual cueing*. If a certain repeated spatial configuration of items is associated with the target appearing at a particular location, observers will respond to the target more quickly when this display is presented, even though they have no conscious awareness of the relationship. Target locations can also be associated with repeated photographs or background colors.

A dramatic example of contextual influences is the *prevalence effect:* Observers are more likely to miss rare targets than frequent targets. This effect occurs for a number of reasons. When discrimination is difficult, it takes more evidence to convince an observer that an item is indeed a rare target. Observers are also more likely to terminate the search early and say "no" when the likelihood of a target is low. Finally, when asked to respond with a "yes" or "no" key press, it can be difficult to suppress the "no" response when one has been hitting that key again and again for hundreds of trials.

Major Theories of Visual Search

The seminal account of visual search is Treisman's *feature integration theory,* which placed visual search experiments at the center of attention research. According to Treisman, the visual system analyzes an image into a set of basic visual features (e.g., color) arranged in independent *feature maps* and computed in parallel. A green target among red distractors pops out because it is the only item producing activation on the green feature map. Focal attention is necessary to bind features from different maps into a single object percept. So finding the green vertical conjunction (or the T) requires a serial process of directing attention to each item in turn until the target is found. Treisman's model has undergone substantial revision over the last quarter-century, but this fundamental insistence on the need for attention to bind features has remained.

A related model is Jeremy Wolfe's *guided search theory,* which proposes that feature information can guide attention to likely targets. *Bottom-up* information indicates which items are more different from their neighbors, while *top-down* information boosts items that share target features. Both types of information are combined into a *salience map* indicating the likelihood that an item is a target without indicating why. In the conjunction search previously described, green items would produce more activation than red items, and vertical items more activation than horizontal items. The most active location on the salience map is probably, but not necessarily, the green vertical target. This explains why conjunction search is intermediate in difficulty between feature and spatial configuration searches.

Feature integration theory and guided search theory are both *early selection* theories in that they assume that attention acts at an early processing stage to combine feature information. John Duncan and Glyn Humphreys proposed a *late selection* theory of visual search, *attentional engagement theory*. This theory is closely allied to the *biased competition* theory of attention later formalized by Duncan and Robert Desimone. The essence of biased competition is that objects compete for access to neural resources at several stages in the visual stream. This competition is biased by attention, such that objects with desired characteristics are more likely to win the competition. In visual search, this means that items that resemble the target are boosted, whereas items that are unlike the target are inhibited. Activation and inhibition spread among items according to their similarity. A field with uniform distractors is easy to search because inhibition spreads easily. When the distractors are heterogeneous, inhibition does not spread as easily, and the search is less efficient. When distractors share features with the target (as in conjunction search), the search becomes more difficult because inhibition of distractors can spread to the target, just as activation can spread from the target to distractors.

The two early selection theories disagree with the attentional engagement theory as to the location of the attentional bottleneck. A newer class of models has attempted to account for visual search without assuming any attentional limitations. These *decision-limited* theories of search, building on signal detection theory, assume that set size effects arise at the decision stage. In some models, each additional item increases the uncertainty in the display. In others, set size effects arise because signals from distractors are pooled with those from targets.

Each of these theories can account for a broad, overlapping range of findings. Feature integration, guided search, and attentional engagement theories have primarily tried to account for reaction time data (particularly search slopes). In contrast, decision-limited theories have been primarily designed to explain accuracy for briefly presented and masked search arrays. A major theoretical project in the field is now to bridge the reaction time and accuracy domains.

Todd S. Horowitz

See also Attention: Object Based; Attention: Spatial; Eye Movements: Behavioral; Feature Integration Theory; Signal Detection Theory and Procedures; Visual Scene Perception

Further Readings

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*(1), 28–71.

Duncan, J., & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review, 96*(3), 433–458.

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & Psychophysics, 62*(3), 425–451.

Rosenholtz, R. (2001). Search asymmetries? What search asymmetries? *Perception & Psychophysics, 63*(3), 476–489.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1–66.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202–238.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5*(6), 495–501.

VISUAL SPATIAL FREQUENCY ANALYSIS

Looking at any visual scene, we seem to effortlessly parse the world into objects and their backgrounds. We don't know exactly how this process is accomplished in the visual system, but we do know that at the earliest stages of pattern perception, the brain analyzes the scene, piece by piece, into a patchwork of local representations. These representations signal how much energy is present at different spatial scales and different orientations at each spatial location. The process through which this is accomplished is *visual spatial frequency analysis*. Presumably, at higher levels of processing, the pieces are put back together to efficiently represent a scene. This entry describes contrast and contrast sensitivity, multiple independent channels and Fourier analysis, and channel interactions.

Contrast and Contrast Sensitivity

A fundamental property of the visual system is that it responds to change, over time or over space. When the change is a change in emitted luminance (as on a TV screen) or a change in the reflectance properties of a surface (light versus dark paint, for example) over space, we see an edge, an object, or a pattern. The difference between the light and dark areas is known as the contrast of a pattern. Contrast can be very high, as black versus white, or it can be very low, as in the difference between two similar gray shades. Our sensitivity to contrast depends considerably on the size, or spatial scale of a pattern, and to a lesser extent upon its orientation. Sinusoidal grating patterns are often used to assess sensitivity at different spatial scales and/or orientations. These are one-dimensional patterns that gradually change over space in luminance, high and low, sinusoidally around some mean light level. They look like fuzzy stripes.

Sinusoidal gratings are defined by four characteristics: contrast, orientation, phase, and spatial frequency. The depth of the excursion—how much the sinusoid deviates from the mean luminance level—is a measure of the contrast of the grating. Every grating also has a single orientation, e.g., vertical, horizontal, or tilted. Each grating has spatial phase, which is the position of the grating relative to some specified reference point. Spatial phase is important because it specifies how, in real objects that have multiple sinusoidal components, these components add together to create the perception of, for example, an edge. The grating's spatial frequency is a measure of

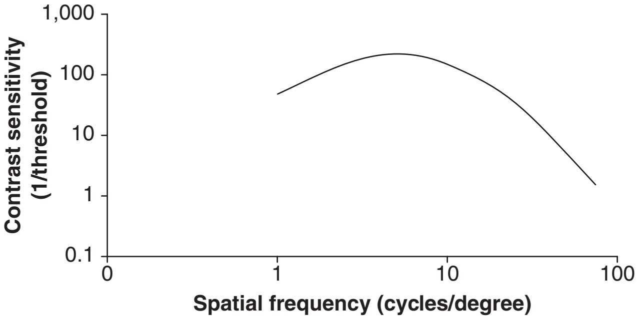


Figure 1 Contrast Sensitivity Function

Notes: A typical contrast sensitivity function showing sensitivity (1/threshold contrast) plotted against spatial frequency. Both axes are in log units.

how many cycles (a dark bar and a light bar) occur in a specified unit of space. A very fine grating, with many cycles, is a high frequency grating. A coarse gating, with few cycles, is a low frequency grating. Because perceived spatial frequency depends upon viewing distance (a fine grating appears coarser at very close distances), spatial frequency is usually expressed in terms of its retinal image size. In particular, it is expressed in cycles per degree of visual angle, often abbreviated c/deg or c^*deg^{-1}. Your thumb at arm's length covers approximately two degrees of visual angle on your retinal surface.

At very low contrasts, a pattern is just barely discernible. How much contrast is needed to just barely perceive that a pattern is present (its threshold contrast) depends primarily upon its spatial frequency. A graph plotting sensitivity (1/threshold contrast) against spatial frequency yields a curve that is typically an inverted U-shape, ranging from about 0.1 c/deg up to about 60 c/deg. Peak sensitivity is at approximately 3–10 c/deg. The curve is known as a contrast sensitivity function (CSF), and reflects not only how well the optics of the eyes transfer light, but also the relative sensitivity of the neural portion of the pattern processing system to different spatial frequencies.

Multiple Independent Channels and Fourier Analysis

For many years, it was thought that the contrast sensitivity function represented a single sensitivity profile of all neurons involved in processing patterns. This was known as the single channel model. In the late 1950s, David Hubel and Torsten Weisel reported what would become Nobel Prize–winning results: They recorded from an early level of processing in cat visual cortex and found cells that were very localized with respect to spatial location in the visual field, responding only to a very restricted area of stimulation. The area to which the cell responds is known as its receptive field. These cells were found to have different preferred orientations and spatial frequencies, responding better to some combinations and less or not at all to others. That is, their receptive fields differed with respect to preferred spatial scale and orientation, refuting the idea that all cells were similar in their sensitivity profiles. These results cast doubt on the single channel model and provided new ways to think about pattern processing. About a decade later, a number of researchers presented the first behavioral (psychophysical) evidence for multiple channels, and several researchers almost simultaneously proposed various formal psychophysical multiple channels models as alternatives to the single channel model. Although differing in detail, these models all proposed that the contrast sensitivity function was actually the sensitivity envelope of a number of more narrowly tuned channels or filters, each selective to a small range of spatial frequencies and orientations. The ranges overlapped from filter to filter, but it was assumed that if spatial frequencies and/or orientations of two sinusoidal components differed sufficiently, entirely separate and independent channels were activated. Thus, the multiple channels model is often referred to as the independent multiple channels model.

Much of the evidence for the multiple channels model is based on behavioral data gathered in psychophysical studies, such as the one by Colin Blakemore and Fergus Campbell, who measured the CSFs of various individuals. They then adapted the visual system by having the individuals view a very high-contrast sinusoidal grating of a particular frequency. The result of adaptation is decreased sensitivity. Blakemore and Campbell then remeasured the individuals' CSFs. They repeated the experiment using different adapting frequencies. The logic of the study was that if the CSF represented a single channel, adaptation would depress sensitivity along the entire curve, and different adapting frequencies would be equally effective. If, however, the CSF was the sensitivity envelope of a set of more narrowly tuned spatial frequency channels, the CSF would be depressed only in the vicinity of the adapting frequency, and the location of the depression would vary with the adapting frequency. The latter result was found and provided strong support for the multiple channels model.

One idea that became widely adopted was that by virtue of monitoring activity in each of the tuned cells, the visual system was performing something loosely akin to a Fourier analysis on the retinal image. Fourier analysis is based on the French physicist Baron Jean Fourier's theorem that all real patterns that vary in time or space, no matter how complex, can be specified by a sum of simpler components, such as sine and cosine waves. For example, a square-wave pattern that alternates black and white stripes can be built up from a series of sine and cosine components that differ in frequency, amplitude, and phase. The higher frequencies modify the lowest spatial frequencies by squaring up their edges and reducing the fuzzy appearance.

Fourier analysis provides the mathematical tools to specify what frequencies are present. Neurons tuned to specific spatial frequencies and orientations might accomplish much the same task by acting as filters, becoming activated only when certain spatial frequencies and orientations are present in a stimulus. The idea of a sensory system performing something like the neural equivalent of a Fourier analysis had earlier been applied to the neural response to complex tones in auditory processing, and now forms the basis for our understanding of the sense of hearing.

An important part of specifying stimuli and responses in the Fourier domain is that the principles of linear systems then apply. One important principle is that the response of a linear system to a compound stimulus, composed, say, of two or three sinusoidal gratings, is simply the sum of its response to each grating separately. This means that if the response of the visual system to simple sinusoids is known, the response to any arbitrary complex pattern can be easily predicted. This proposition was widely tested by many detection experiments using low contrast stimuli, and in almost all cases, it was found to hold to a first approximation.

An important problem with this notion in the visual domain, however, is that both cell responses and psychophysical performance are nonlinear as contrast increases. They have a threshold below which no response is seen, and while responses increase linearly as contrast increases over a limited range, they saturate, or stop increasing, at some contrast level. The models were developed to incorporate these nonlinearities as separate and distinct processes from the linear filtering stages of processing. In this way, the principles of linear systems could continue to be used to predict responses to complex patterns, but were modified by the within-channel nonlinearities.

Channel Interactions

More recently, the notion of independent channels for different orientations and spatial frequencies has been called into question. For example, using high-contrast stimuli, both psychophysical performance and individual cells have been shown to be indirectly affected by stimuli whose orientations are orthogonal to a test component—far enough away to be part of a different spatial frequency channel. In the case of individual cells, the cell does not respond directly to the orthogonal component, but when activated by a stimulus it is sensitive to, the orthogonal component modifies the cell's response by reducing it a considerable amount. This indicates that there must be connections between the cells.

In the psychophysical domain, performance on some fine discrimination tasks has also been shown to be affected by spatial frequencies and orientations that are far apart. Recent models have proposed the existence of one or more sensitivity-regulating pools of neurons that span a wide range of spatial frequencies and orientations. These are known as contrast gain control pools, and reflect between-channel interactions. They add a nonlinear stage to the processing of spatial information.

Many current models in visual processing take the form of a linear-nonlinear-linear processing cascade. The higher-level linear mechanism is often some sort of combining circuit that takes input from selected multiple tuned channels and either compares them or signals some sort of average value. In the context of spatial frequency analysis, one organizing principle along which these circuits operate is that they appear specialized to signal a particular type of information—for example, orientation information or spatial frequency information, but not both. These schemes provide a way to separately encode different stimulus dimensions at intermediate stages of processing, even though they are jointly represented in the first stages of cortical processing. These are the intermediate steps in ultimately creating an organized, meaningful percept of the world.

Lynn A. Olzak

See also Contrast Perception; Spatial Layout Perception, Neural; Visual Processing: Primary Visual Cortex; Visual Stimuli

Further Readings

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology, 203,* 237–260.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197,* 551–566.

DeValois, R. L., & DeValois, K. K. (1990). *Spatial vision.* New York: Oxford University Press.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America. A, Optics and Image Science, 11*(6), 1710–1719.

Graham, N. V. S. (1989). *Visual pattern analyzers.* New York: Oxford University Press.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9*(2), 181–197.

Olzak, L. A., & Thomas, J. P. (1986). Seeing spatial patterns. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1). New York: John Wiley & Sons.

Olzak, L. A., & Thomas, J. P. (1999). Neural recoding in human pattern vision: Model and mechanisms. *Vision Research, 39*(2), 231–256.

VISUAL STIMULI

A *stimulus* is usually thought of as some physical event, or set of events, that is likely to lead to a sensory, perceptual, or behavioral response. The concept has its roots in sensory physiology, in the study of sensation and perception, and in behavioral psychology. Although it seems straightforward to consider a flash of light to be a stimulus for vision, the concept is more difficult than it seems. For example, the flash of light at its source (the distal stimulus) and the light that reaches the cornea and then the retina of the eye (the proximal stimulus) are different and need to be distinguished. Some theorists, notably J. J. Gibson, have argued that the concept should be abandoned altogether on the grounds that perceptual systems are attuned to the *information* that is embedded in the complex patterns of energy captured by the sense organs and are not prodded or goaded into responses by "stimuli." Nevertheless, the term remains in wide use and this entry considers the variety of often ingenious "stimuli" that have been devised to reveal the capabilities of the visual system (e.g., visual acuity, contrast sensitivity, detection of motion) and the nature of the underlying mechanisms.

Historically, studies in visual physiology and psychophysics have tended to use fairly simple stimuli that could be accurately measured and controlled (spots or bars of light, simple shapes, stripes, checkerboards), combined with simple tasks and simple responses ("did you see the light?" "were the stripes vertical or horizontal?" "was A brighter than B?"). The choice of stimulus depends on the aspect of vision being tested and the type of question being asked. In the 19th century, visual stimuli were limited to what could be produced by pen and ink, light sources, mechanical, optical, and photographic systems, but the development of electronics and then computer displays and virtual reality in the 20th century has led to a rich and varied range of possibilities. In all cases, measurement and calibration of the display device, and knowledge of its limitations, are crucial aspects of quantitative research.

Spatial Vision

A small *spot* of light (Figure 1a) may be considered an elementary stimulus because it is localized in space and stimulates a small patch of the retina. A *sinusoidal grating* (Figure 1c), on the other hand, stimulates a large region of visual (and retinal) space. Considered as a wave pattern oscillating across space, however, it is quite simple because the wavelength and the orientation of the pattern are constant. The *spatial frequency* of the grating is given by the reciprocal of the wavelength: high spatial frequency gratings have narrow bars, whereas low frequencies have wide, blurred bars. *Fourier analysis* is a way of describing complex patterns (such as sound waves) as a combination of pure sine waves of different frequencies added together. The same can be done for visual images, but one needs to specify both the frequency and orientation of the sine wave components, as in Figure 1(c). Thinking of visual images both in spatial terms (as a collection of spots or pixels) and in Fourier terms (as a collection of sine wave components) has become important in modern optics and vision research. As a compromise between these extremes, the *Gabor patch* (a sine wave grating limited by a smooth "window" or "envelope"; Figure 1b) has been widely used because it is localized to some degree in visual space and has a well-defined frequency and orientation. The Gabor

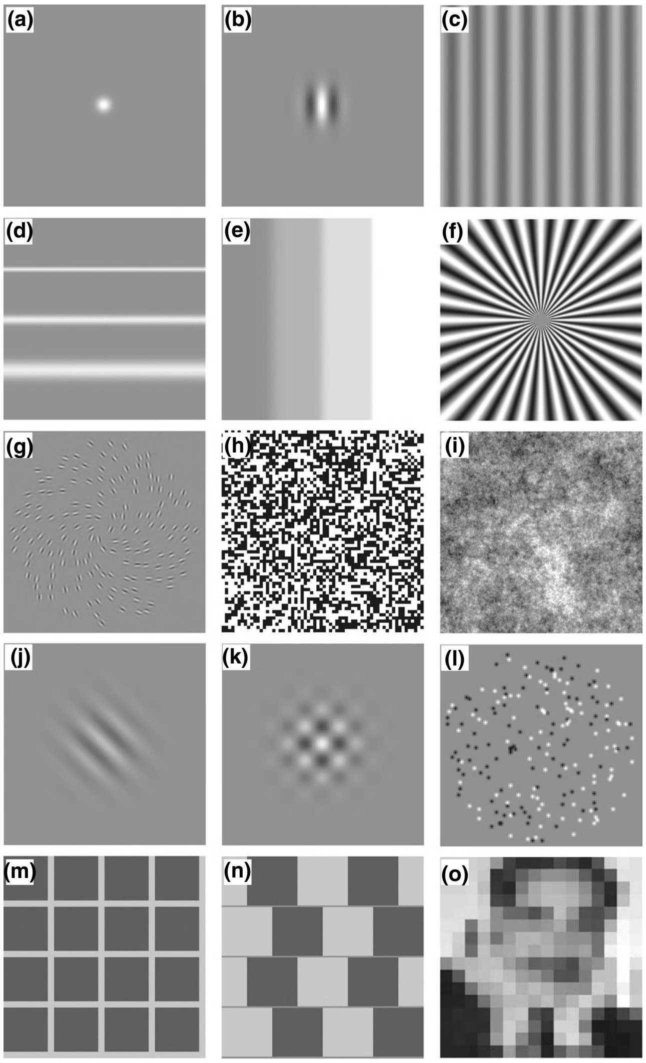


Figure 1 A Variety of Images Used as Stimuli in Vision Research

patch is thought to resemble the receptive fields of simple cells in the visual cortex, and thus to stimulate them fairly selectively in visual experiments. Important parameters defining these patches are the size and location of the envelope, the spatial frequency and orientation of the grating, as well as its position ("phase") relative to the center of the envelope, and its *contrast.* Single lines of different widths (Figure 1d) and edges of different blurs (Figure 1e) are also useful elementary stimuli. More complex sine wave images, such as the radial grating (Figure 1f) and spiral or concentric versions of this, display all spatial frequencies and orientations at the same time and can evoke vivid visual aftereffects.

Texture, Contours, and Depth Perception

Repeated or randomly scattered elements tend to create visual regions with a distinct *texture* whose appearance depends on the elements and their arrangement. Natural textures include sand, wood grain, bark, fur, and gravel, whereas synthetic textures can be created by repetition of simple elements such as dots, bars, or Gabor patches (Figure 1g). Textures can often be rapidly recognized and distinguished, and may represent a more global, large-scale "stream" of visual processing, distinct from that which enables us to identify the local elements. Spatial variations in texture (e.g., Figure 1g) contribute to the recognition of 3-D object shape and to the separation (segmentation) of an object from its background or from other objects. Contour perception is adept at picking out smooth curves, even across disconnected elements (Figure 1g).

Surprisingly, even random patterns are useful research tools. Random "noise" (Figure 1h) has been widely used in the study and clinical assessment of stereoscopic depth perception because it eliminates all cues to shape or form, except the slight difference (binocular disparity) between images shown to the left and right eyes. Noise also has a distinct texture that changes when the noise is filtered to change the spatial frequency content of the image. Figure 1(i) shows random noise filtered to reduce the strength of its high spatial frequencies. It has a spatial frequency range that is similar to natural images and a cloudlike appearance because, with random phases, it has no coherent spatial structure.

Motion

Motion perception has been studied with drifting, expanding, or rotating versions of all these 1-D and 2-D images. An expanding image often appears to move toward or away from the observer, and if it is large enough, it may make observers feel that they are moving, or even being turned upside down, as in some fairground illusions and cinematic special effects. Moving plaids (Figure 1k), the sum of right and left oblique components (Figure 1j), and arrays of dots moving in partially random directions (Figure 11) have been especially important in revealing the sequence of local and more global processing stages in visual motion analysis.

Special-Purpose and Natural Images

In addition to these widely used generic images, many images are used to study specific phenomena, such as the illusory dark spots at the intersections in the Herman grid (Figure 1m) or the strong distortions of orientation seen in the Café Wall or Munsterberg illusion (Figure 1n). Block-sampled images (Figure 1o) have been used in research and used by the media and the police to disguise identity; but not always successfully, because if you blur your eyes or view from a distance, Groucho Marx should emerge from the apparently random squares. Thus, natural images such as faces, houses, objects, and scenes, often modified in some way, are gradually bridging the gap between studies of the initial stages of visual image analysis and the later stages of visual perception and recognition.

Mark Georgeson

See also Contrast Perception; Light Measurement; Motion Perception; Visual Illusions; Visual Scene Perception; Visual Scene Statistics; Visual Spatial Frequency Analysis

Further Readings

Bruce, V., Green, P. R., & Georgeson, M. A. (2003). *Visual perception: physiology, psychology and ecology*. Hove, UK and New York: Psychology Press.

Carpenter, R. H. S., & Robson, J. G. (1999). *Vision research: A practical guide to laboratory methods.* Oxford, UK: Oxford University Press.

Wandell, B. A. (1995). *Foundations of vision.* Sunderland, MA: Sinauer.

VISUAL SYSTEM: EVOLUTION OF

Visual systems include light-detecting cells (photoreceptors), typically eyes as organs with image forming capacity, and neuron networks that extract useful information from the patterns of light that activate the photoreceptors. This information is used in numerous ways, guiding daily and annual cycles of activity, mating, migration, and other behaviors, as well as motor actions that preserve life and lead to successful reproduction. Eyes have proven to be so useful that they have evolved at least twice and probably many times. They vary in many ways, while consistently providing information about wavelength and intensity of light and sharing molecular mechanisms of photo detection. Thus, some of the same genes are used in the development of independently evolved photoreceptors and eyes.

Even though light detection does not depend on having eyes, eyes are structures that protect receptors from damage when focusing or channeling light so that at least a crude image of the external scene is displayed across a sheet of receptors. The invertebrate ancestors of the first vertebrates already had a bilateral body plan with a head, and an eye on each side of the head. The eye was a chambered or camera type of eye that used a lens to focus an image onto the receptor sheet (retina). Invertebrate cephalopods (squid and octopus) have independently evolved the same type of eye, although the details of construction are much different. Yet, these similar eyes allow some vertebrates and cephalopods to have excellent vision. In both lines of evolution, useful vision depends not only on the type of eye and its evolutionary modifications but also on the neural networks that extract information from the eye. The vision of invertebrates is an interesting topic because they have many types of eyes and photoreceptive organs, but the focus here is on the visual system of vertebrates, with an emphasis on the mammalian visual system, especially the visual systems of primates.

The visual system has been modified in many ways in vertebrate evolution. A dramatic change took place with the advent of mammals, as the primitive neocortex proved to be a place where the visual system could be greatly expanded and altered. This entry only touches on some of the modifications that have occurred in the visual systems of mammals, and the major changes have clearly taken place in the visual systems of primates. To a large extent, the impressive visual abilities of most primates relate to the great expansion of the visual cortex, and the subdivisions of the visual cortex into functionally distinct areas, and areas into patterns of modules. However, changes in the visual cortex were also accompanied by widespread modifications in other parts of the visual system. The visual pulvinar expanded and became divided into more nuclei. The superior colliculus became more dependent on cortical inputs, and its retinotopic organization was altered to match that of the visual cortex. The retina became dominated by the parvocellular (P) class of ganglion cells. These P cells projected almost exclusively to the dorsal lateral geniculate nucleus and provided information for form and color vision. The dorsal lateral geniculate nucleus, subdivided into three types of layers, corresponded to three types of input from the retina. Many more alterations have been described, and more are likely to be discovered. While the primate visual system has unique specializations, each of the major branches of the primate radiation has its own specializations of the visual system.

The Basic Visual System

The first vertebrates were fish. Present-day fish constitute a highly diverse group, including the highly successful ray-fined fish, as well as the lungfish, whose ancestors gave rise to the first tetrapods. Comparative studies of fish and other vertebrates indicate that early vertebrates already had a visual system with a retina that projected largely or exclusively to contralateral targets on the forebrain and midbrain. The retina included the photoreceptor cells, the rods and cones, as the deepest layer. The basic functions of rods for dim light and cones for bright light and color have been conserved across vertebrate species. Rods typically dominate, and in most nonmammalian vertebrates, rods and cones activate a common second-order neuron, the bipolar cell that in turn activates the ganglion cell that provides the output of the retina. Horizontal cells and amacrine cells function as interneurons in the retina, with horizontal cells contacting each other, and other horizontal cells and amacrine cells are activated by bipolar cells and provide inhibitory output to ganglion cells. Retinal projections included those to the neuropil of the contralateral thalamus in the region of the dorsal and ventral lateral geniculate nuclei, the pretectum, and the optic tectum (the superior colliculus of mammals). Retinal projections also terminated in the suprachiasmatic nucleus of the hypothalamus where visual information is used to synchronize behavior with daily and seasonal changes in the day–night light cycle. Thus, major subcortical components of the human visual system are very ancient.

The major target of the retina of early vertebrates, the optic tectum, had descending outputs that terminated largely contralaterally in the brain stem and upper spinal cord, where they had access to motor neuron groups controlling eye and body movements so that objects of interest could be better visualized. Ascending outputs of the optic tectum were to the pretectum, a thalamic region that may be a homolog of the mammalian visual pulvinar, and the region of the dorsal and ventral lateral geniculate nuclei. The pretectal and accessory optic nuclei, via connections with motor nuclei, were involved in oculomotor and other visual reflexes. The visual information that reached the dorsal thalamus probably influenced a number of visually guided behaviors.

The Emergence of the Mammalian Visual System

In present-day reptiles (and probably in the reptilian ancestors of mammals), the dorsal lateral geniculate projects to the dorsal cortex, the section of the forebrain that became the neocortex in mammals. However, the dorsal cortex of reptiles is a small, thin structure, and it does not have the major role in vision that the visual cortex has in mammals. Although the visual system of the ancestors of mammals already contained a small cortical component, the greatly expanded visual cortex of mammals is highly involved in object recognition and the visual guidance of behavior.

Instead of a small, thin sheet of dorsal cortex, mammals have a thick, proportionally large sheet of neocortex that is traditionally divided into six layers, each specialized for different functions. In addition, the cortex is divided into cortical areas, sometimes called the organs of the brain. These cortical areas have different functions, such as visual, auditory, somatosensory, motor, planning, emotion, integration, and other functions. Much of the story of the evolution of the mammalian visual system has to do with the neocortex, as variations in the organization of the neocortex, and often, the visual cortex, have permitted mammals to adapt to many different environmental conditions. Of course, transformations in any part of the visual system mediate transformations in other parts. In particular, the expansion and differentiation of the visual neocortex that occurred in some mammals required changes in the visual information that was delivered to the cortex, and alterations took place in the retina, dorsal lateral geniculate nucleus, superior colliculus (optic tectum), and pulvinar of the visual thalamus.

In order to get more information to the visual cortex, the mammalian dorsal lateral geniculate nucleus is larger than those in reptiles, and it has new features of internal organization. In all mammals there are inputs from the ipsilateral as well as the contralateral eye, and these two types of input are segregated from each other. Typically, the ipsilateral input is much smaller, and it is separated into a small layer in the middle of a mass of neurons innervated by the contralateral eye. In primates, three classes of retinal ganglia cells innervate separate layers in the lateral geniculate nucleus. In other mammals, the dorsal lateral geniculate nucleus is organized in various other ways. The various types of organization probably reflect altered capacities of the visual system as the visual cortex expanded and changed in different ways.

The other source of visual information to the cortex depends on the superior colliculus (optic tectum), which projects to several locations in the posterior thalamus. In reptiles, this projection zone is called nucleus rotundus, but in mammals this thalamic region is called the pulvinar or the pulvinar complex. The parts of the pulvinar with inputs from the superior colliculus project to areas of the visual cortex, providing a secondary source of visual information. In primates, and several other highly visual taxa, the visual pulvinar is greatly enlarged and subdivided into six or more nuclei. Only some of these nuclei get input from the superior colliculus, but all get input from cortical visual areas, while projecting back to cortical visual areas, further distributing visual information.

The superior colliculus has also been altered by the development of mammalian visual cortex, as the superior colliculus receives projections from areas of the visual cortex, as well as direct retinal inputs. In primates and some other mammals, ipsilateral retinal inputs have greatly increased in magnitude. In primates, something even more remarkable has happened. Most mammals have a primitive pattern of retinotopic organization (representation) in each colliculus so that the entire retina of the contralateral eye is represented. However, all mammals have a different type of representation in the visual cortex, where visual areas of each cerebral hemisphere represent inputs from only the parts of the retinas of the two eyes that image the contralateral half of the visual field (the contralateral hemifield). Thus, areas of the visual cortex of each cerebral hemisphere represent the contralateral visual hemifield, not the whole field of the contralateral eye. This type of representation provides a substrate for binocular vision. The hemifield type of representation is also found in the superior colliculus of primates, but not other mammals. Thus, both old and new parts of visual systems can be modified in evolution, often together as they are interrelated.

The Visual Cortex in Primates and Other Mammals

Almost all mammals have at least two areas of visual cortex, the primary area (visual area 1 or V1) and the secondary area (V2). V1 receives activating inputs from the dorsal lateral geniculate nucleus, and V2 receives activating inputs from V1. V1 also projects to a medially adjoining area, sometimes called prostriata, which is poorly understood and has uncertain functions. In early mammals, and many present-day mammals, V1 and V2 project broadly to other nonvisual areas to provide visual information. Both V1 and V2 contain an orderly retinotopic representation of the contralateral visual field and are considered important areas for object recognition. Mammals with small brains and little neocortex typically have 3 to 5 visual areas. Some clearly have more: cats and other carnivores, for example. A number of other mammalian taxa have not been adequately investigated. But primates have a greatly altered visual system with many visual areas. Most notably, only primates have the middle temporal visual area (MT) in the upper temporal lobe. MT is a visual area that is specialized for detecting visual motion. MT is surrounded by a constellation of four associated visual areas that are also involved in processing visual motion signals, and these areas are also unique to primates. The cortical projections of MT and associated visual areas are to visuo-motor areas of the posterior parietal cortex, where visual information is further processed in order to guide eye movements, reaching for food, and other useful movements. These visuomotor areas add 4 to 8 additional areas to the primate visual system. Other added visual areas are in the temporal lobe, where information from V1 and V2 is processed further in a series of areas devoted to object recognition. Both temporal lobe areas and parietal lobe areas send processed visual information to the frontal lobe, where it directly or indirectly impacts on frontal lobe motor and premotor areas. Overall, the visual cortex of Old World macaque monkeys appears to contain about 35 visual areas. New World monkeys and prosimian primates have many of these areas, but not all, and humans appear to have additional visual areas, such as the fusiform face area that is important in recognizing faces. Thus, the ancestors of early primates had already evolved a complex visual system with a substantial number of new visual areas, those now found in all primates. But the process did not stop, as further areas were added in anthropoid monkeys and in humans. In addition, older visual areas, V1 and V2, have been modified in primates by becoming modular in organization. Both areas have been systematically subdivided into smaller functional units. V2, for example, has been divided into alternating bands of three types, differing in inputs, outputs, and neuron response properties. Thus, V2 of primates functions as three visual areas. V1 has regions where neurons are grouped according to how they respond to lines and bars of different orientations (orientation selective modules) and smaller foci of high metabolic activity (measured with cytochrome oxidase) where neurons are less involved with stimulus orientation. Other visual areas, such as MT, also appear to be subdivided into two or more functional classes of modules.

Jon Kaas

See also Retinal Anatomy; Vision; Visual Processing: Extrastriate Cortex; Visual Processing: Primary Visual Cortex; Visual Processing: Retinal; Visual System Structure

Further Readings

Casagrande, V. A., Khaytin, I., & Boyd, J. (2007). The evolution of parallel visual pathways in the brains of primates. In J. H. Kaas & T. M. Preuss (Eds.), *Evolution of nervous systems, Volume 4, Primates* (pp. 87–108). London: Elsevier.

Kaas, J. H. (2003). The evolution of the visual system in primates. In L. Chalupa & J. D. Werner (Eds.), *The visual neurosciences,* Cambridge: MIT Press.

Lyon D. C. (2007). The evolution of visual cortex and visual systems. In J. H. Kaas & L. A. Krubitzer (Eds.), *Evolution of nervous systems, Volume 3, Mammals* (pp. 267–306). London: Elsevier.

VISUAL SYSTEM STRUCTURE

Humans and other animals that rely principally on their sense of sight to perceive their environment have highly evolved visual systems. The visual system consists of billions of connections within and between many visual areas in the brain. The information that traverses these connections operates almost entirely beneath levels of consciousness, such that individuals can observe, plan, and act within their environments with extreme precision. This entry describes feedforward and feedback flow, parallel processing streams, and the visual system structure.

Feedforward and Feedback Flow

Upon first examination, the visual system appears to be organized in a serial fashion. In other words, “lower” visual brain areas connect to “higher” areas in a hierarchical progression. The term *feedforward* relates to this idea that visual information is fed forward from a lower brain area to the next, higher one. While there is much structural and physiological evidence to support this idea, it would be a mistake to characterize visual processing as merely a feedforward process. Instead, it is becoming increasingly clear that as visual information flows serially from one visual area to the next, there is also a constant backward, or *feedback* flow of information. This combination of feedforward and feedback flow allows visual information to move bidirectionally, allowing for different areas of the visual brain to engage in ongoing communication with one another.

Parallel Processing Streams

Visual information that flows between brain areas is also separated into distinct pathways or “streams.” These streams carry different kinds of information about the visual environment. They also run parallel to one another, traveling from one visual area to the next. At particular brain areas, information is exchanged or shared between different streams. As visual information flows to higher brain areas, the quality of the information being carried in each stream changes to reflect this exchange of information. The visual system utilizes this parallel processing strategy to save time and space. Parallel streams allow information to travel simultaneously along separate pathways, increasing the overall rate of information processing. They also underlie a common infrastructure that can decode a wide variety of visual stimuli.

The visual system of humans and many primates begins with three parallel processing streams, analogous to three separate cables carrying separate information. These streams are the magnocellular (M), parvocellular (P), and koniocellular (K) streams. Each stream is characterized by structurally distinct components, the neurons, and by physiologically distinct properties. Traditionally, the M-stream is thought to carry information about the motion of visual stimuli; the P-stream is thought to carry information about the finer details of visual stimuli, including some color information; and although much less is known about the K-stream, it is thought to carry information about additional color properties of visual stimuli. The M-, P-, and K-streams remain separated from their point of origin in the retina to the first visual cortical area (V1), where their information is first exchanged. As visual information progresses to higher visual cortical areas, from the occipital cortex to parietal and temporal visual areas, the distinctions between M-, P-, and K-streams become less pronounced. Instead, two new streams emerge: The dorsal stream, referred to as the "where" pathway, which carries motion signals and is thus highly influenced by the M-stream; and the ventral stream, or "what" pathway, which carries information about the details of stimuli, including color and shape, and is traditionally thought to be more influenced by the P-stream.

Visual System Structure

The Eye

The eye is a complicated structure designed to focus light onto a layer of receptor neurons. These receptors, called photoreceptors, are located within the retina, which is a multilayered carpet of neurons situated on the rear, interior surface of the eyeball. The job of photoreceptors is to convert, or transduce, the light signal into a neural signal (i.e., a language that is comprehensible to neurons).

The retinas of humans and many other species contain two families of photoreceptors (see figure 29, color insert). Rod photoreceptors make up the majority of photoreceptors in the primate retina and are distributed rather evenly across the retina except at the center, or fovea, where their density decreases. Rods are best at detecting photons in low light conditions, like nighttime. Cones make up a second family of photo-receptors and are concentrated in the foveal region of the retina. Cones are responsible for vision in conditions of bright light (i.e., daytime). There are three distinct types of cone photoreceptor: those sensitive to long (L), medium (M) and short (S) wavelengths of light. Humans and other animals that have three cone types (L, M, and S) are referred to as trichromats and are capable of trichromatic color vision. Other animals, including some primates, dogs, and cats, have only two types of cone photoreceptors and therefore see the world as dichromats.

Photoreceptors communicate changes in light levels to bipolar neurons located in the inner layer of the retina. Bipolar cells are On or Off, depending on whether they respond to increments (On) or decrements (Off) of light. In addition to being On or Off, bipolar cells are also segregated into at least three different neuronal types based on their association with the magno-, parvo-, or koniocellular parallel processing streams. Diffuse bipolar neurons of the M-stream receive inputs from all three cone types. For this reason, the M-pathway is not as well suited for carrying information about color as it is for carrying information about luminance, which is related to the perceptual experience of brightness. Diffuse bipolar neurons connect to parasol retinal ganglion cells, which carry M-stream information out of the retina and on to the next structure in the chain of communication: the visual thalamus.

The P-stream begins with midget bipolar neurons, which receive information from individual L or M cones in the foveal region of the retina. Midget bipolar neurons then project to midget retinal ganglion cells, which also carry their P-stream information to the visual thalamus. Because foveal midget bipolar and retinal ganglion cells keep their L and M inputs segregated through the retina, the P-stream is traditionally described as carrying L/M color information.

Less is known about the neurons of the K-stream; however, it appears that S-cone bipolar neurons receive their inputs specifically from S-cones. S-cone bipolar neurons synapse onto small bistratified retinal ganglion cells, which project to the visual thalamus. Because of the S-cone-specific nature of K-stream neurons in the retina, the K-stream is thought to carry S-channel color information to the visual cortex (color insert Figure 35).

Retinal bipolar and ganglion cells reflect some sophisticated physiological properties. This is because the concept of multidirectional information flow is already at play in the retina. While information flows "forward" from photoreceptors to bipolar neurons to retinal ganglion cells, there is also "lateral" communication between neighboring bipolar cells and retinal ganglion cells, mediated by a large variety of interneurons that display many different anatomical and physiological characteristics.

In order to describe the physiology of retinal neurons, the important concept of the receptive field must be introduced. Every visual neuron has a receptive field, which is the area of visual space in which some change in a visual stimulus will evoke a response in the neuron. Receptive fields are smallest in the foveal region of the retina and become larger with distance away from the fovea. For this reason, our visual acuity is greatest at the fovea. Retinal neurons respond best when particular patterns of light are presented within their receptive field: a light spot on a dark background or a dark spot on a light background. This type of receptive field structure is called center-surround in reference to the opposite responses evoked in the center and surround of the neuron's receptive field. M-stream retinal neurons have centers and surrounds that are opposite in luminance (i.e., light/dark). Foveal P-stream retinal neurons have centers and surrounds that are L/M opponent.

The Visual Thalamus

The retina sends axons to many targets in the brain, but the target most important for communicating information to the cerebral cortex for perception is the lateral geniculate nucleus (LGN) of the thalamus. The LGN in humans and many monkey species contains 6 principal layers and 6 intercalated, or "in-between" layers. LGN layers 1 and 2 are the magnocellular layers, which receive M-stream input from the contralateral (or opposite side) and ipsilateral (or same-side) retinas respectively. Layers 3 through 6 are the parvocellular layers which receive P-stream input from alternating ipsilateral and contralateral eyes. The 6 intercalated layers, also called koniocellular layers, receive K-stream input, also from alternating eyes.

The axons of retinal ganglion cells synapse directly onto LGN neurons that then project to the first visual cortical area, V1. Communication between retinal ganglion cells and LGN neurons is described as "driving" because the physiological properties of LGN neurons closely resemble those of their retinal inputs. For example, LGN neurons of each stream display similar center/surround receptive fields to their retinal ganglion cell inputs. However, the LGN is not simply a relay structure. It contains its own population of interneurons that modulate the visual signal received from the retina via local thalamic circuitry. Additionally, the LGN receives excitatory feedback input from corticogeniculate neurons located in V1. Corticogeniculate neurons synapse directly onto LGN neurons and also communicate indirectly with LGN neurons via an intermediary structure called the reticular nucleus of the thalamus (RTN). The RTN literally wraps around the LGN and sends inhibitory signals to LGN neurons. Thus, the circuit connecting the visual cortex back to the LGN contains both an excitatory signal and an inhibitory signal. These signals are more modulatory in nature; that is, "backward" flowing information does not dramatically change the receptive field structure of LGN neurons. However, corticogeniculate feedback is likely to play an important role in visual information processing by mediating bidirectional communication between the LGN and V1.

The Primary Visual Cortex

LGN neurons send their information to the first visual cortical area, called the primary visual cortex, striate cortex, or V1. Like all of the cerebral cortex, V1 is a laminar structure containing 6 main layers. Feedforward inputs from the LGN provide layer-specific inputs to V1. M-steam LGN neurons project to layer 4Cα with a collateral projection to the bottom portion of layer 6. P-stream LGN neurons target layer 4Cβ and send collateral axons to the top portion of layer 6. K-stream LGN neurons synapse in a variety of V1 layers, depending on their LGN layer of origin. However, many K-stream axons target columnar structures in layers 2/3 termed "blobs" or "patches" in reference to their appearance in tissues stained for cytochrome oxydase (color insert Figure 36).

V1 is the first site where information from the M, P, and K parallel processing streams is mixed. It is also the site where information from the two eyes is combined. This mixing occurs through a myriad of cortical circuits connecting V1 neurons within and between cortical layers. V1 circuits connect a diverse population of both excitatory and inhibitory neurons. Inhibitory neurons all restrict their axonal projections "locally" within V1, often within single cortical columns, and demonstrate large anatomical variation. Excitatory neurons also include a large diversity of anatomically distinct pyramidal and spiny stellate cell types. While all excitatory neurons give rise to local projections within V1, a large proportion also send axonal projections out of V1. One major class of "projecting" neurons targets the thalamus and includes the corticogeniculate neurons previously described. A second major class of projecting neurons target "higher" visual cortical areas, such as V2 and the middle temporal visual area MT, and form the foundations of the dorsal and ventral parallel processing streams.

The pattern of LGN inputs to V1, along with the specificity of neuronal circuits within V1, lie at the heart of an exquisite functional architecture. First, V1 is organized in a visuo-topic (also called retinotopic) fashion that reflects the visual layout of the world. Accordingly, there is a gradual progression across the surface of V1 that corresponds to the progression across visual space. Second, V1 neurons located within the same column spanning the six cortical layers have receptive fields located in the same region of visual space. Not only do neurons in the same cortical column respond to the same region of visual space, they also tend to share other response properties. For example, neurons in the same cortical column typically share similar ocular dominance (eye preference) properties and similar tuning for stimulus orientation. Indeed, ocular dominance and orientation tuning and possibly color preference are mapped in systematic fashion across V1.

The visual response properties of V1 neurons reflect the exchange of information across the M-, P-, and K-streams. V1 neurons typically do not display the elementary center/surround receptive field structure described for retinal and LGN neurons. Instead, many new receptive field properties emerge, including simple and complex receptive fields that display preferences for the orientation, direction, speed of movement, size, and often color of visual stimuli.

The diversity of receptive field types encountered in V1 reflects the vast amount of local information processing that occurs within V1, mediated by local circuits between neurons in different layers and/or columns. Although this field of systems neuroscience is not yet well understood, a number of circuit patterns have been identified that may be ubiquitous across cortical areas and even different species. These patterns represent a generalized flow of information across the cortical layers within a column. As previously described, most geniculocortical information enters V1 in the sublamina of layer 4C. Neurons in layer 4C project their axons mainly up to the more superficial layers (layers 2/3–4B) and down to layers 5 and 6. Layer 6 neurons appear to loop information back to layer 4C, either directly or indirectly via corticogeniculate projections back to the LGN. Neurons in the superficial layers participate in a reciprocal loop with neurons in layer 5. The superficial layers (those above layer 4C) and layer 5 contain the majority of extrastriate projecting neurons allowing V1 to communicate in a feedforward fashion with downstream cortical areas.

Extrastriate Cortical Areas

There are 2 extrastriate cortical areas that receive substantial input from V1: areas V2 and MT. Similar to V1, V2 contains a detailed columnar structure; however, V2's columnar organization differs from that of V1. V2 appears to be organized into "stripes" that appear thick, thin, and pale in tissue stained for cytochrome oxidase (see figure 1 in Action and Vision, p. 7, and figure 30 in color insert). Neurons in the thick stripes communicate in a feedforward fashion with cortical area MT, and neurons in the thin and pale stripes provide feedforward input to cortical area V4. The projection from V2 to MT is part of the dorsal stream or "where" pathway and carries information about stimulus motion, while the projection from V2 to V4 is part of the ventral stream, or "what" pathway and carries information about stimulus color and form. V2 neurons located in all three stripes also provide feedback projections to V1.

Cortical area MT is an integral member of the dorsal stream and contains neurons that are strongly selective for stimulus motion. MT neurons tend to have large receptive fields, indicating they are not well suited for carrying information about the exact location or fine qualitative detail of a visual stimulus. Instead, they are well adapted for detecting changes in the movement of visual stimuli. It is likely that neurons in MT receive V1 and V2 inputs that are dominated by the M-stream with less involvement from the P- and K-streams. Although less is known about the anatomical organization of MT, evidence suggests a large-scale columnar organization. MT neurons provide feedforward input to a variety of cortical areas associated with the dorsal stream, including the middle superior temporal visual area (MST) and areas in the parietal lobe. MT neurons also provide feedback projections to V1 and V2.

Recently, much interest has been focused on the neural processes and cortical areas involved in planning and executing actions. Cortical areas involved in these types of cognitive function need to incorporate information from many different sensory systems (including vision) and work in close association with areas involved with executive function (decision making), as well as motor areas (movement planning). Cortical areas in the parietal regions and inferotemporal (IT) regions of the brain are associated with vision and with planning and executing. These represent some of the furthest reaches of the dorsal and ventral streams, respectively. Neurons recorded in dorsal extrastriate areas such as the lateral intraparietal area (LIP) have been implicated in responding preferentially during planning phases of visually guided tasks, for example. Ventral stream neurons in different IT cortical regions have been identified as preferentially recognizing specific facial features and expressions. Little is known about the cellular and circuit structure underlying these cortical areas; however, more is known about the anatomical projections to and from different extrastriate cortical areas. As a general principle, feedforward and feedback projections become increasingly widespread, contacting a greater number of areas with each progression stage. Interestingly, there appears to be a consistent relationship between extrastriate cortical areas and various thalamic structures, whereby some proportion of outputs from one cortical area to the next may be relayed through these thalamic nuclei.

There is a strong, albeit less well-defined, relationship between visual sensory areas and areas in the frontal lobe, including the prefrontal cortex and the frontal eye fields. Neurons in the prefrontal cortex appear to be involved in coordinating eye movements that have clear significance to visual processing. The prefrontal cortex receives input from many visual areas in the brain, including the extrastriate cortical areas and the superior colliculus. Prefrontal areas involved in eye movements therefore maintain ongoing communication with visual structures at many processing stages.

Farran Briggs and W. Martin Usrey

See also Cortical Organization; Retinal Anatomy; Visual Processing: Extrastriate Cortex; Visual Processing: Primary Visual Cortex; Visual Processing: Subcortical Mechanisms for Gaze Control

Further Readings

Callaway, E. M. (2005). Structure and function of parallel pathways in the primate early visual system. *Journal of Physiology, 566,* 13–19.

Field, G. D., & Chichilnisky, E. J. (2007). Information processing in the primate retina: Circuitry and coding. *Annual Review of Neuroscience, 30,* 1–30.

Maunsell, J. H. R., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience, 10,* 363–401.

Sincich, L. C., & Horton, J. C. (2005). The circuitry of V1 and V2: Integration of color, form and motion. *Annual Review of Neuroscience, 28,* 303–326.

VOMERONASAL SYSTEM

Some animals have sensory systems that humans lack completely, making it difficult for us to even imagine what their perceptual worlds are like. For example, many fish can detect electric fields in the water around them, including, in some cases, electric fields produced by other members of their own species. Other animals have sense organs that detect stimuli outside the ranges that humans can detect, such as insects that can see ultraviolet or polarized light or bats that can hear extremely high frequencies that we call “ultrasound.” The *vomeronasal system,* or accessory olfactory system, is a sensory system lacking in humans, but present in most tetrapods (amphibians, reptiles, and mammals). Although our understanding of the sense of smell has undergone a revolution in the past two decades, the vomeronasal system is still so poorly understood that we don't even know if it functions like an extra sense for animals that have it, or simply extends the range of odorants that can be detected by their noses.

The vomeronasal system is physically separate from the rest of the olfactory system: The two systems have discrete sense organs, the sensory epithelia and receptor neurons inside the organs are different, and the axons of the receptor neurons project to distinct regions of the forebrain. In mammals, the vomeronasal organs are a pair of cigar-shaped organs located between the floor of the nasal cavity and the roof of the mouth. In reptiles, they are often called Jacobson's organs and vary considerably in size, shape, and location; for example, in most snakes the vomeronasal organs are roughly spherical, large, and are located above the roof of the mouth. In contrast, in some turtles the vomeronasal organs are little more than grooves in the floor of the nasal cavity. In salamanders, the vomeronasal organs are generally pouches that extend from the outer edge of the nasal cavity; and in frogs, the organs are usually separate from and below the nasal cavities.

Overall, vomeronasal receptor neurons look like other olfactory receptor neurons, except that the tips of the dendrites, where the cells contact odorants, are covered in tiny hairs called microvilli. The receptor neurons in the olfactory epithelium are more diverse, and their dendrites can be covered in larger hairlike cilia, or microvilli, or both. Both the receptor molecules and the ion channels that open when an odorant is detected differ completely between vomeronasal and other olfactory receptor neurons. These disparities in appearance and biochemistry suggest that vomeronasal receptor neurons detect different odorants, or respond differently to odorants, than do other olfactory receptor neurons.

The receptor neurons in the vomeronasal organ have long axons that project to the accessory olfactory bulb in the forebrain. The accessory olfactory bulb is generally located behind the main olfactory bulb, which receives input from the receptor neurons in the olfactory epithelium. The types of neurons and circuits they form differ between the accessory and main olfactory bulbs, again suggesting that the vomeronasal system performs a different function than the rest of the olfactory system.

The main and accessory and olfactory bulbs project to distinct regions in the forebrain. The main olfactory bulb has extensive projections to the cortex and hippocampus, as well as to portions of the amygdala, whereas the accessory olfactory bulb projects mainly to the amygdala. The different regions of the amygdala that receive input from the main and accessory olfactory bulbs are interconnected and are the first place in the brain where vomeronasal and other olfactory input converge. Because the vomeronasal system makes no direct connections with the cortex but the main olfactory system does, some have suggested that stimulation of the vomeronasal system may not give rise to conscious perception of odorants, whereas stimulation of the olfactory epithelium might.

Although a distinct vomeronasal system is present in most tetrapods, it has been lost many times over the course of vertebrate evolution. The system has been lost in the family of aquatic salamanders containing mudpuppies, in crocodilians and birds, and in whales and dolphins. Within bats alone, the vomeronasal system has been lost a dozen or more times; it has also been lost in Old World primates, including humans. In addition, recent research indicates that elements of the vomeronasal system are present in fish. The types of odorant receptors and ion channels that characterize vomeronasal neurons in mammals are present in sensory neurons in the nose of bony fish, such as goldfish, pufferfish, and zebrafish, as well as possibly lampreys. In zebrafish, these neurons have been shown to send their axons to a region of the olfactory bulb that is separate from the regions receiving input from olfactory-type neurons. These results suggest that vomeronasal neurons may have existed, intermingled with other olfactory neurons, in the noses of the earliest vertebrates. For reasons we don't yet understand, these neurons then became a separate sensory organ in the fishy ancestors of tetrapods.

Interestingly, animals that possess vomeronasal organs often have specialized adaptations for getting odorants inside the organ. The clearest example comes from snakes, which pick up molecules when they flick their forked tongues and then deposit the molecules in their vomeronasal organs. In some mammals, such as rodents, the entire organ contracts periodically, pumping in molecules that are dissolved in the mucus surrounding the opening to the organ. Cats and hoofed mammals make a facial grimace called "flehmen" that is believed to facilitate entry of molecules into the vomeronasal organ.

These adaptations may provide clues to the function of the vomeronasal system, which remains obscure. One relatively early idea was that the vomeronasal system might be specialized for processing pheromones. (Pheromones are chemicals released by animals that elicit specific behavioral or physiological reactions from other members of the same species, often in the context of reproduction.) Some insects, such as moths, have sensory hairs and brain regions that are specialized for processing pheromonal information, and researchers thought that vertebrates might also have such specializations. In addition, it might make sense for animals to have a general, flexible sensory system for detecting any odorant that might come along, and a separate, "hard-wired" sensory system for responding in a precise way to pheromones. The main olfactory system was thus thought to be flexible and general, and the vomeronasal system hard-wired for mediating responses to pheromones. This general dichotomy fits some of the data. For example, male mice experience a testosterone surge and begin to produce an ultrasonic courtship call when they smell urine from a reproductive female, and this effect is due entirely to vomeronasal sampling of the urine. On the other hand, many studies have shown that the vomeronasal system also responds to compounds that are not pheromones—for example, snakes use vomeronasal cues to track prey. In addition, some pheromones work by activating the main olfactory system. The best-studied vertebrate pheromone is androstenone, which is produced by male pigs, and its effects on females are mediated entirely through the main olfactory system. Overall, the data demonstrate that the vomeronasal system is not specialized for pheromone detection.

Another hypothesis is that the vomeronasal organ detects large molecules, which are too heavy to float high enough in the air to contact the receptor neurons in the olfactory epithelium. Instead, these odorants might gain entry to the vomeronasal organ by being picked up on a snake's tongue or drawn into the organ as it pumps. If this idea is correct, it would explain why the vomeronasal system seems to have been lost in reptiles and mammals, such as dolphins and bats, that live their lives in water or up in the air: They would not encounter heavy molecules, and therefore would have no need for a specialized sensory system to detect them. This hypothesis might also explain why the vomeronasal organ often seems to be the sense organ for pheromones—it happens that many pheromones, such as sodefrin, a newt pheromone, are large, heavy molecules. Pheromones that activate the main olfactory system, like androstenone, are small molecules that can easily become airborne. Thus, the "molecular size" hypothesis seems to be a good explanation of the function of the vomeronasal system. A few difficulties need to be resolved, however. First, we now know that fish have both vomeronasal and regular olfactory receptor neurons in their noses, and molecular size or ability to become airborne is irrelevant for animals that live in water. Another problem is that some heavy particles, such as viruses, are known to reach the olfactory epithelium, and surely small, light molecules can gain access to the vomeronasal organ. Until we better understand the relationship between the chemical structure of odorants and the molecular receptors that they bind to, it will be difficult to determine whether the molecular size hypothesis is correct or not.

Heather L. Eisthen

See also Consciousness; Olfaction; Olfaction: Evolution of; Olfaction and Reproductive Behavior; Olfactory Bulb: Functional Architecture; Olfactory Central Processing; Olfactory Receptors and Transduction; Olfactory Stimulus; Pheromones

Further Readings

Baxi, K. N., Dorries, K. M., & Eisthen, H. L. (2006). Is the vomeronasal system really specialized for detecting pheromones? *Trends in Neurosciences, 29,* 1–7.

Halpern, M., & Martinez-Marcos, A. (2003). Structure and function of the vomeronasal system: An update. *Progress in Neurobiology, 70,* 245–318.

Meredith, M. (2001). Human vomeronasal organ: A critical review of best and worst cases. *Chemical Senses, 26,* 433–445.

Restrepo, D., Arellano, J., Oliva, A. M., Schaefer, M. L., & Lin, W. (2004). Emerging views on the distinct but related roles of the main and accessory olfactory systems in responsiveness to chemosensory signals in mice. *Hormones and Behavior, 46,* 247–256.

W

WEIGHT PERCEPTION

In popular language, *weight* can mean an object (such as a paperweight), the weight of that object, or the mass of that object. In physics, a distinction is made between *weight* and *mass*. Weight is a force, usually the result of Earth's gravity acting on an object; weight is defined as mass times acceleration, so weight varies under different gravitational or accelerative forces. Mass is the amount of matter in an object, and it can be measured by the ratio of an applied force to the resulting acceleration; it remains constant under different forces. On Earth, weight information is normally available through the pressure sense because of the constant acceleration of gravity. Inertial mass information is also available if observers move the object and note the force they have applied and the resulting acceleration. Under weightless conditions, as in orbital space travel, only mass information is available. In this entry, *weight/mass* is used for those situations where both weight and mass information is available. *Heaviness* is the subjective feeling of weight or mass; there is no reported difference in sensation relating to the two physical concepts.

We pick up and handle objects for various purposes. Sometimes the purpose is to judge the weight/mass and decide whether a package is too heavy for the mail or whether one tool is heavier than another. We are not very accurate, so we use weighing scales for precision. More often the purpose is to move an object from one place to another. We look at the object, plan how to grip it and the likely force required to move it, and then carry out the action while making rapid corrections if necessary. We obtain feedback from the muscle, pressure, and joint position senses, and the whole operation gives us a sense of heaviness. This entry discusses our ability to discriminate between masses, how sensory magnitude increases with mass, the causes of illusions and other errors, and adaptation to altered force environments.

Weight/Mass Discrimination

Weight/mass perception was among the earliest of the senses studied by experimental psychologists. Ernst Heinrich Weber—famous for stating Weber's law in 1834—used weight/mass as an example of his law: the smallest difference in weight/mass that can be discriminated is proportional to the mass of the objects. Thus, if 105 grams can be distinguished from 100 grams, 210 grams would be needed to detect a difference from 200 grams. This gives a Weber fraction of 5/100 or 0.05. Weber was interested in sensory mechanisms. He measured weight discrimination with touch alone (on a supported hand) and weight/mass discrimination with touch plus the muscular sense (with lifting), and found the latter much superior. He concluded that the "weight sense" depended both on touch and on the muscle sense. Modern experiments show the typical fraction for touch alone to be about 0.13, and that for lifted objects to be around 0.09, (though the best values for well-practiced observers may be as low as 0.02). As with other sense modalities, the fraction is not constant at all intensities but is high for very light objects, low for medium objects, and rises again for very heavy objects. Discrimination is best when the observer is well adapted to the range of masses in the test and is worse when the range varies.

Sensory Magnitude

Many authors have investigated how the sensation of heaviness increases with the mass of the lifted object. Gustav Fechner (a contemporary of Weber) believed that sensory magnitude was a logarithmic function of the stimulus value, increasing more slowly than the stimulus intensity. He argued this on mathematical grounds (deriving it from assumptions about Weber's law) and did not directly measure sensory magnitude. S. S. Stevens (working in the mid-20th century) believed that sensation was a power function of the stimulus value. He used direct scaling methods in which observers assigned numbers to their sensations, and he found an upwardly rising curve with sensory magnitude increasing faster than the stimulus value. In practice, many different curves have been produced, depending on the test method. One serious flaw in magnitude experiments where the test objects are visible is that the containers are all the same size, to avoid size being a cue to mass. The result is that the estimates are affected by the size-weight illusion, in which the absence of a change in size makes the heavier and denser containers feel extra heavy.

Weight/Mass Illusions

There are several well-known weight/mass illusions, such as the size-weight illusion (first described by Charpentier in 1891), the material-weight illusion, and the color-weight illusion. The size-weight illusion has the largest effect, with perceived ratios of up to 1.5 between two equal weights. These illusions can be described as contrast with the expected weight/mass. If two objects are equal in mass but different in size, the larger one feels lighter when lifted; if they are equal in mass and volume but differ in surface appearance, the one that appears to be made of the denser material feels lighter. The effect of color or brightness is similar to that of material: the darker color feels lighter, presumably because darker colors are associated with denser materials. These expectations are evident if observers are asked to say which looks heavier without lifting the objects. There are other examples of expectations. Observers can be trained with a heavy object on the left and a lighter one on the right; if they are then given two equal objects, the one on the right feels heavier.

To say that these are all examples of contrast with the expected weight/mass does not suggest a mechanism. A more precise idea is that there is a mismatch between expected and actual sensory feedback, owing to the manner of lifting. If an object is expected to be heavy, it is lifted with more force than one expected to be light; it then rises more easily than expected, causing it to feel light. On this account, the observer is unaware of the actual force he has applied, but is aware of the feedback. Attempts to measure the manner of lifting have led to mixed results. Some experiments show that the illusion persists even though observers scale their fingertip forces to the true weights of objects. Observers reduce their reliance on visual size cues and learn to apply the appropriate force, but the heaviness illusion remains. This seems to refute the sensorimotor mismatch hypothesis, while allowing for a cognitive mismatch. Perhaps cognitive predictions are independent of sensorimotor predictions, involving different parts of the brain. One theory is that sensorimotor processing takes place in the dorsal stream (parietal lobe), whereas the heaviness illusion experience takes place in the ventral stream (temporal lobe). Weight illusions also occur when objects are placed on a static supported hand, which shows that the manner of lifting need not be involved. Changes in the manner of lifting may be a side effect rather than a cause of the illusion, or they may be one of several contributing factors.

Mass Constancy and the Manner of Lifting

The illusions mentioned previously are all examples of the misperception of weight/mass. Nevertheless, we try to judge the mass of objects rather than the force required to move or lift them. We obtain prior information about an object's likely mass from its visual appearance, and we adjust our manner of lifting appropriately. Objects feel almost equally heavy when lifted by different parts of the body, despite the fact that different muscle groups and degrees of effort are involved. More effort is required to lift an object with an extended arm than to lift it close to the body. Similarly, objects suspended at different lengths along a lever differ little in heaviness, provided the observer can see the situation. Observers can also compensate for different weights added to their own arm.

However, mass constancy is imperfect. Heaviness depends partly on perceived effort due to maintaining grip, so slippery objects feel heavier than rough objects. This suggests that observers fail to distinguish adequately between grip force and lifting force and are influenced by the total effort. Objects feel lighter when grasped with a wide grip than a narrow grip because a wide grip requires less effort. Objects also feel lighter when they have a larger contact area or are being lifted by more fingers. Objects lifted by two hands feel slightly lighter than when lifted by one hand, so there is some integration between the two hands.

Altered Force Environments

We normally lift objects against the force of Earth's gravity (1 g, equivalent to 9.8 m/s^2). This force acts through the z-axis (main body axis) of an upright human. High g forces can be produced in the human centrifuge, and brief periods of both high and low g are evident in parabolic flight, whereas spaceflight offers prolonged microgravity (near 0 g). The effective weight varies in altered force environments, doubling under a 2 g force and becoming 0 under microgravity. Water also alters the forces because the upthrust of the water reduces the object's weight, while more effort is required to overcome viscous drag.

Mass Constancy

Swimmers partially adapt to the weight reduction in water; they judge objects in water as heavier than their effective weight and show an aftereffect in air when they judge objects as heavier than before they entered the water. Experiments at different g levels in the centrifuge and parabolic flight show that heaviness estimates are midway between the effective weight and the weight at 1 g, revealing a tendency to mass constancy. Observers probably gain knowledge of the force environment through the increase or decrease in the weight of their own arms and body.

Weight/Mass Discrimination

Weight/mass discrimination is impaired in altered force environments, though it improves with adaptation, and aftereffects of impaired discrimination are usually shown on return to 1 g. These findings apply to objects in water, in the centrifuge, in parabolic flight, and in spaceflight. Microgravity has a more harmful effect than high force environments because it removes the pressure and kinesthetic information that is always present when observers hold objects against Earth's gravity. Under microgravity astronauts must shake or move objects and note both the force they have imparted and the resulting acceleration. Videos of astronauts' hand movements when judging mass show that they impart higher peak accelerations (shorter rapid movements) under microgravity than under 1 g, perhaps because this improves discrimination. Changes in hand movements should not be a source of error, provided astronauts correctly monitor their movements. However, analysis of postflight videos on Earth suggests that hand movements can cause errors. Incorrect weight/mass judgments were correlated with unusually large or small movements. Astronauts probably interpret a large hand movement (low peak acceleration) as due to a light object, and a small movement (high peak acceleration) as due to a heavy object, although the variation lay in their own production rather than in the mass of the object. This finding is consistent with those size-weight illusion studies that implicate unmonitored changes in the manner of lifting, but is inconsistent with those where the illusion is independent of the manner of lifting.

Helen E. Ross

See also Haptics; Perceptual-Motor Integration; Proprioception; Reaching and Grasping

Further Readings

Bock, O. (1998). Problems of sensorimotor coordination in weightlessness. *Brain Research Reviews, 28,* 155–160.

Flanagan, J. R., & Beltzner, M. A. (2000). Independence of perceptual and sensorimotor predictions in the size-weight illusion. *Nature Neuroscience, 3,* 737–741.

Jones, L. A. (1986). Perception of force and weight: Theory and research. *Psychological Bulletin, 100,* 29–42.

McIntosh, R. D. (2000). Seeing size and weight. *Trends in Cognitive Science, 4,* 442–444.

Murray, D. J., Ellis, R. R., Bandomir, C. A., & Ross, H. E. (1999). Charpentier (1891) on the size-weight illusion. *Perception and Psychophysics, 61,* 1681–1685.

Ross, H. E. (1969). When is a weight not illusory? *Quarterly Journal of Experimental Psychology, 21,* 346–355.

Ross, H. E., Schwartz, E., & Emmerson, P. (1987). The nature of sensorimotor adaptation to altered G-levels: Evidence from mass discrimination. *Aviation Space and Environmental Medicine, 58*(Suppl. 9), A148–A152.

WINE TASTING

There are probably as many perceptions about wine tasting as there are tasters. For some, it is a charade; it can be an expression of cultural heritage; it may increase self-esteem; but primarily, it is a means of enhancing wine appreciation. Most people taste wine holistically and strictly on a subjective basis. Wine competitions are more objective, ranking wine against some standard—the grape variety, geographic region, production style, or artistic attributes. The latter include features termed complexity, harmony, subtlety, dynamism, development, duration, and uniqueness. The most objective form of tasting is descriptive analysis. In it, panelists are trained and selected for their use of specific sensory terms. These are used to describe the wine analytically. All tasting forms have their particular purpose and merits. However, their different intentions mean that their conclusions typically differ. In addition, any tasting possesses a degree of subjectivity due to differences in taster experience and sensory acuity. This entry describes the various factors that influence assessment of the taste of wine.

Detailed tasting follows a particular sequence. This involves visual assessment, followed by smelling, then tasting. The process is completed by assessing overall quality. The more detailed the tasting, the more individual attributes are described as to how they change in character and intensity.

Most consumers taste wine with food. At its best, this is viewed as pairing particular wines with specific dishes. However, the combination frequently results in mutual flavor reduction, rather than enhancement. For example, when sampled together, cheese appears milder and wine less sour or bitter/astringent. The ambiance of a tasting can also significantly influence the impression. A regional wine so pleasant when sampled in a local bistro on a relaxed vacation may be disappointing back home. In addition, the supposed synchrony between local wines and cuisines reflects more habituation than a conscious effort to harmonize regional produce.

For the general populace, critical tasting is typically viewed as ranking wines. As noted, this is usually based on wine being assessed relative to established norms. These criteria are, however, not strictly objective, and can change with the social context and have varied over time. Even within categories, subtypes exist, for example, the eight quality categories of German wines. In addition, various countries may have distinctive expressions of varietal wines. An excellent example is found with Sauvignon Blanc. New Zealand versions are the most flavorful and varietally unique, whereas those from Australia, California, France, and South Africa are less varietally distinct but still show regional attributes.

Although the sensory acuity of experienced tasters has only occasionally been investigated, exceptional skill has not been demonstrated. Even sommeliers may have difficulty consistently detecting the regional or varietal origin of wine. In addition, identifying similar wines, sampled minutes apart, is a severe challenge. This should not be surprising; the differences can be subtle. In addition, they often vary from year to year, region to region, and producer to producer. What training and experience create is a memory bank of features that characterize different groups of wines. Although training and experience improve judging skill, it also modifies qualitative judgment. For example, trained tasters are often critical of features consumers may not detect or may consider of little importance. Thus, expert opinion usually differs considerably from that of most wine consumers.

Frequently, wine flavors are described in terms of fruits, flowers, vegetables, and so forth. Although wines do possess compounds similar to those in plants, most analogies to fruit and the like bear faint resemblance to the real thing. It is not without reason that most wine descriptions tell more about the taster than the wine (their past experiences or emotional reaction to the wine). Exceptions are the bell pepper odor of some Cabernet wines and the lychee aspect of some Gewürztraminer wines. These varietal fragrances are generated by the same compounds that donate the characteristic flavors to the respective fruits. Because human perception is highly susceptible to suggestion, mention of a particular term can induce others to think they detect the same feature, especially when they experience difficulty in describing the wine's aroma.

As noted, contextual factors can significantly influence wine perception. Studies have shown that white wines (colored red) are described using terms typically applied to red wines, not white wines. In addition, an ordinary wine is far more likely to be described glowingly if thought to be of high quality than if its true identity is known.

Because fragrant compounds give wines their most distinctive attributes, they feature prominently in critical tastings. Taste and mouthfeel aspects are primarily involved in assessing overall flavor quality. However, for those little accustomed to wine, taste is the principal aspect influencing preference. Many consumers gloss over the intricacies of wine fragrance, so critical to the connoisseur and professional taster. The features that distinguish fine wines are widely acknowledged by professionals, but little known to most consumers.

The most critical type of tasting (descriptive analysis) is a laboratory tool used to determine if modifications in grape cultivation or wine production detectably affect a wine's characteristics and to what degree. The technique may also help determine if geographic regions produce distinctively different wines. Whether such subtle differences can be detectable, or are important, is a moot point to most consumers. Nevertheless, descriptive analysis is critical to product development and an unbiased sensory assessment of a wine's attributes.

Consumers are most aware of tastings reported in food and wine magazines and in newspaper columns. Their practical use can be assessed only by sampling some of the wines oneself. If the writer's perception resembles one's own, their recommendations can be useful in directing wine purchases. More generally, reports in the popular press have their principal benefit in encouraging consumers to expand their wine selection and explore their own perceptive preferences.

Ronald S. Jackson

See also Flavor; Olfaction; Taste

Further Readings

Broadbent, M. (2003). *Michael Broadbent's wine tasting.* London: Mitchell Beasley.

Jackson, R. S. (2009). *Wine tasting. A professional handbook* (2nd ed.). London: Academic Press.

Lehrer, A. (1983). *Wine and conversation.* Bloomington: Indiana University Press.

WORD RECOGNITION

Humans have the remarkable capacity to express their thoughts, beliefs, and intentions through a physical medium (sounds, gestures, or pictograms) to share with others. Language's expressive power resides in the ability to express any novel thought by combining elements (e.g., cow, purple, ticklish) into a sequence that conveys that thought (e.g., a ticklish purple cow). Thus, a critical component of language comprehension consists of recognizing the presence of these elements, the words, in the speaker's discourse in order to retrieve their individual meanings and combine them to derive an overall interpretation of the discourse. Research on word recognition has focused on understanding how people categorize a physical (auditory or visual) token as one of the many words they know. This entry first reviews the factors that affect the perceptual choice inherent to the recognition of spoken or printed words. It then discusses how people represent the forms of words they know in order to recognize them in speech or print. Finally, the entry briefly discusses how people's linguistic knowledge influences their evaluation of sensory information.

Word Recognition as a Perceptual Choice

People recognize the words of their language remarkably quickly and accurately, whether written or spoken.

However, this ease conceals the complexity of the process. The roughly 50,000 words that an adult knows are formed by combining sound or letter elements drawn from a much smaller set (e.g., 26 letters and between 40 and 45 distinct sounds in English, depending on the dialect). As a result, words resemble one another to a large degree. The physical difference between the words *cat* and *cot* is subtle but critical to attend to if the correct meaning is to be retrieved. Thus, recognition entails selecting the best matching hypothesis among alternatives.

Importantly, these alternatives exert an influence on the outcome. Under some conditions, words are harder to recognize if they have many "neighbors," that is, if they resemble many other words. For instance, while reading, people spend more time fixating on printed words with many neighbors than words with few, suggesting that the former take longer to identify. Likewise, spoken words that sound similar to other words are recognized more slowly, and less accurately when presented in noise, than spoken words with fewer neighbors. These findings are often taken as indicating that recognition is a competitive process, where the spoken or written stimulus is compared to every hypothesis simultaneously and where the best matching hypothesis is identified with respect to the overall support accrued by the alternatives. However, under some conditions, similarity to many other words facilitates a word's recognition. For example, written words that share all but one of their letters with many other words (e.g., *cave,* similar to at least 16 other words such as *cake, have, gave, care*) are recognized faster than words with fewer such neighbors (e.g., *next,* similar to only 4 other words, *newt, neat, nest,* and *text*). That neighbors can facilitate and impede a word's recognition has been taken as reflecting the existence of two separate mechanisms by which neighbors affect a word's recognition. Neighbors provide mutually incompatible interpretations of the same portion of the stimulus, the part that differs among them. The more alternatives, the harder the perceptual choice becomes. But neighbors also share some elements, and the perception of these parts is enhanced by the support provided by their occurrence in multiple words. Whether neighbors facilitate or impede the overall outcome depends on which influence outweighs the other.

The ease of recognizing a word is also affected by how often it has been encountered before, with frequent words being recognized more accurately and faster than rarer words. In studying this effect, frequency of occurrence is estimated by the number of times each word occurs in a large sample of spoken and written language. The so-called *word-frequency effect* has sometimes been explained as reflecting the ease with which an operation is performed given the number of times it has been performed in the past. This account is supported by the neural mechanism that underlies learning, whereby frequent exposures to a given stimulus strengthen the connections between the population of neurons that encode the physical attributes of the stimulus and those that encode its interpretation. However, the word-frequency effect can also be attributed to the influence of biases (or priors) on perception. This view assumes that the physical attributes of a word take time to apprehend and decode and provide only probabilistic cues, and people are generally making a decision under some uncertainty. Under these conditions, an optimal strategy for deciding whether or not a given word is present in the stimulus is to consider the probability of encountering the word at all. Frequent words have a high base rate, and people's guesses reflect this rate. Frequent words are recognized more accurately than rare words because, based on noisy information from the stimulus, the best bet is to assume that the word is a frequent one. Rare words are recognized more slowly or less accurately because, initially at least, people favor their more frequent alternatives.

Finally, the process of recognizing spoken and printed words differs because of differences in their medium. A written word is a sequence of elements spatially organized, and the reader controls the uptake of sensory information. By contrast, a spoken word is a sequence of elements temporally organized, and the order and the speed at which sensory information is received is not under the listener's control. Speech is a complex, transient, and rapidly changing signal. Because humans' sensory memory is limited, speech must be evaluated and interpreted incrementally rather than, say, phrase by phrase. Real-time processing is at the core of William Marslen-Wilson's seminal work on the recognition of spoken words. According to Marslen-Wilson's "cohort" model, the first sounds of a word determine a cohort of hypotheses that are compatible with this early information. Subsequent information serves to prune the hypotheses that are no longer supported by the signal. The point in time at which a spoken word is recognized can be precisely identified as the point where the input has excluded all candidates but one (i.e., the word's *uniqueness point*). Although Marslen-Wilson revised his theory to allow phonetic input to provide gradient, rather than all-or-none, support to word candidates, the assumption that word recognition makes immediate use of the signal as it unfolds has prevailed.

How Are Word Forms Represented?

In order to decide which word they see or hear, people compare sensory information to an internal representation of the forms that words take. How should this internal representation be characterized? Although there is no definitive answer to this question, issues pertaining to the processing of visual or auditory information have helped identify the central properties that internal representation must possess. Here, some of these issues are reviewed.

How do people judge which word they see or hear based on the sensory information available? One possibility is that they compare the visual or auditory stimulus to each auditory or visual abstract image or template associated with the words they know and determine which template provides the best match. However, considering the number of words and their similarity, templates would be difficult to tell apart. An alternative consists of identifying components of words, which represent a much smaller set, and establish their arrangement. This information distinguishes any word from its alternatives.

In the domain of written words, where most of the research has been conducted on English by presenting printed as opposed to handwritten words, each letter is a symbol that is neatly separated from the surrounding letters and reproduced identically across words (although its image on the retina changes as a function of its size, spatial location, and other low-level factors). Thus, letters are discrete elements that are juxtaposed to form words. Letter recognition seems a necessary component of printed word recognition. The situation is quite different in the auditory domain. The articulation of a speech sound is greatly influenced by the sounds that precede and follow it. For example, the shape and position of the lips prior to articulating the sound *b* change dramatically as a function of the vowel that follows it, as in *bat* or *boot*. This contextual influence on the articulation of speech sounds is reflected in their acoustic realization. It has been proven difficult to identify acoustic properties that instances of the same speech sound produced across contexts all share. Consequently, whether the recognition of a spoken word involves the identification of its subcomponents is a highly debated question. The perception of cursive handwriting has not been studied much but raises similar issues as speech perception.

Because elements are shared among many different words, establishing their order is critical to correct identification. The letter sequence "top" should be interpreted as the word *top* and not *opt*, even though *opt* contains the same elements. Thus, the comparison between sensory information and word-form representations requires an alignment point. For the recognition of visual words, the spaces surrounding a string of letters offer such alignment points. Each letter in the stimulus is viewed as occupying a given slot in space whose position is established with respect to the beginning of the sequence (the left space). The match between the stimulus and a given word hypothesis is evaluated by comparing each letter slot independently. Accordingly, *opt* and *top* have no aligned letters and cannot be confused. However, this view of alignment fails to account for readers' ability to recognize some words with transposed letters (e.g., "jugde" for "judge") as easily as their intact counterparts or the strongly perceived similarity between words like *pluck* and *luck*, which, under the left-aligned scheme, share no overlap.

The alignment issue is even more salient in speech because the alignment between the sensory information and particular word hypotheses must be assessed on a temporal, as opposed to spatial, dimension. For example, the sensory information that results from the phrase *the catalog* is compatible with the words *cat, a, log, cattle,* and *catalogue,* among others. (Rules of syntax and semantics, as well as fine-grained acoustic details, help disambiguate the utterance's interpretation.) However, these word hypotheses entertain different relationships with one another, and establishing the relationship among these word candidates requires a precise alignment with the stimulus. *Cat, cattle,* and *catalog* are mutually incompatible hypotheses because they account for some of the same portion of the input, roughly the syllable “cat.” Said slightly differently, the syllable “cat” in speech may be attributed to the word *cat, cattle,* or *catalog,* but to only one of them. Likewise, *catalog* and *log* are mutually incompatible hypotheses because they are different interpretations of the same portion of the signal, the syllable “log.” By contrast, *cat* and *log* are mutually compatible because they each account for a different part of the stimulus. Thus, the perceptual system must represent the temporal dimension at a fine-grained level.

How Does Prior Knowledge Influence Evaluation of Sensory Information?

Word recognition relies on assessing a perceptual stimulus in the light of prior knowledge—the words of one’s language. The interdependence between sensory information and prior knowledge was demonstrated in one of the earliest empirical studies of word recognition, conducted by James Cattell at the end of the 19th century. Under tachistoscopic conditions, where written stimuli are presented for a brief period of time, Cattell showed that letters were identified more accurately when presented in words than in random letter sequences. This finding was later extended by Gerald Reicher, who showed that individual letters were identified faster when occurring in a word context than in isolation. This latter result is especially puzzling if one takes letter recognition as a prerequisite to word recognition. The so-called *word-superiority effect* has been extensively studied, with demonstrations that it extends to the auditory modality. For example, listeners tend to interpret a sound that is ambiguous between the sounds “g” and “k” as “g” if followed by “ift” and as “k” if followed by “iss.” Two competing accounts for the effect have been offered. According to one account, knowledge of which letter sequences form words can penetrate the mechanisms of perception and thereby enhance, or even reshape, the perception of the letters. This view was implemented by James McClelland and David Rumelhart in a cascaded and interactive model of word recognition. According to these models, sensory information provides initial support to some letter hypotheses, which in turn provide support for the words that contain them; these hypotheses then feed back their support to their letter constituents. This process takes place very rapidly and before a single percept has been isolated. Thus, letter recognition can affect and be affected by word recognition until the process reaches a stable state. Importantly, the feedback from words to letters reflects the influence of knowledge on perception. An alternative account of the word-superiority effect offered by Dominic Massaro and colleagues maintains independence between the sources of information. The sensory information associated with each letter is integrated with sensory information extracted from the context, that is, the other letters. The intersection of these two sources of information constrains the set of possible letters at each position in the word more than each source independently. Identification reflects the integration of both sources of information without assuming that one affects the other. Distinguishing these two competing accounts empirically has proved difficult.

Delphine Dahan

See also Eye Movements and Reading; Reading Typography; Speech Perception; Top-Down and Bottom-Up Processing

Further Readings

Cattell, J. M. (1886). The time taken up by cerebral operations. *Mind, 11,* 220–242.

Dahan, D., & Magnuson, J. S. (2006). Spoken-word recognition. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 249–283). Amsterdam: Elsevier.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25,* 71–102.

Massaro, D. W., & Cohen, M. M. (1991). Integration versus interactive activation: The joint influence of stimulus and context in perception. *Cognitive Psychology, 23,* 558–614.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

Rastle, K. (2007). Visual word recognition. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 71–87). Oxford, UK: Oxford University Press.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81,* 274–280.

Index

Note: Entry titles are in **bold.**